RECENZII DESPRE CEI NOUĂ TITANI TECH

„*Cei nouă titani tech* e o carte provocatoare, lizibilă și în care te regăsești cu ușurință. Amy Webb demonstrează o cunoaştere aprofundată a științei din spatele inteligenţei artificiale și a tensiunilor geopolitice care ar putea apărea, mai ales între China și Statele Unite. Autoarea ne oferă o analiză profundă a felului în care inteligenţa artificială ne-ar putea recalibra economiile și actuala ordine mondială, oferind un plan amănunţit prin care omenirea ar putea să-și aleagă un viitor mai bun”.

— Anja Manuel, Universitatea Stanford, co-fondator și partener la Rice Hadley Gates

„*Cei nouă titani tech* e o carte importantă și cu idei proaspete care ne pune în perspectivă potenţialul și pericolele inteligenţei artificiale. Îşi va sluji inteligenţa artificială stăpânii din Washington, Silicon Valley și Wall Street, sau va sluji ea interesele publicului larg? Va reuși ea să concentreze sau să disperseze puterea economică și geopolitică? Graţie lui Amy Webb putem înţelege întrebările și cum putem ajunge la răspunsurile care vor fi mai bune pentru omenire decât parcursul pe care ne aflăm deocamdată. *Cei nouă titani tech* ar trebui introdusă în sălile de clasă și de consiliu din lumea întreagă”.

— Alee Ross, autorul *Industries of the Future*

„*Cei nouă titani tech* ne oferă perspective îndrăzneţe referitor la viitorul inteligenţei artificiale. Spre deosebire însă de mulţi alţii care încearcă să anticipeze viitorul, Webb nu se folosește de senzaţionalism ci de argumente studiate, un context istoric profund și un nivel înfricoșător de plauzibilitate”.

— Jonathan Zittrain, profesor de drept internaţional și informatică, Universitatea Harvard „*Cei nouă titani tech* ne pune pe gânduri, privind situaţia în perspectivă și mai ales punând întrebările corecte referitoarea la inteligenţa artificială și oferindu-ne o foaie de parcurs pentru un viitor optimist împreună cu inteligenţa artificială”.

— Peter Schwartz, autorul cărţii *The Art of the Long*

View

„Cei nouă titani tech oferă argumente importante pentru a refuza abordările orientate spre prezent și a nu considera că firmele care dezvoltă inteligenţa artificială sunt morale prin definiţie. Scenariile propuse de Webb pentru viitor sunt superbe, oferind o viziune amănunţită atât asupra aspectelor care trebuie evitate cât și a celor care trebuie îndeplinite”.

— John C. Havens, director general IEEE Global Iniţiative on Ethics of Autonomous and Intelligent Sistems, autor al cărţii *Heartificial Intelligence: Embracing our Humanity to Maximize Machines*

CEI NOUĂ TITANI TECH

CEI NOUĂ TITAN ITECH

Cum va schimba inteligenţa artificială cursul omenirii

AMY WEBB

The original title of this book is:

The Big Nine: how the tech titans and their thinking machines could warp humanity by Amy Webb

Copyright © 22019 by Amy Webb

Copyright © Editura Globo, 2019, pentru ediţia în limba română

Descrierea CIP a Bibliotecii Naţionale a României WEBB, AMY

Cei 9 titani tech: cum va schimba inteligenţa artificială cursul omenirii / Amy Webb.

— Bucureşti: Globo, 2019

ISBN 978 – 606 – 9653 – 00 – 5

004

Traducător: Răzvan Lerescu Redactor: Diana Morăraşu Tehnoredactor: Sorin Sava

Comenzi: telefon: 0726228911 www.editura globo.ro comenziedituraglobo.ro *Tatălui meu, Don Webb, omul cel mai inteligent pe care l-am cunoscut vreodată.*

CONŢINUT

Introducere: înainte de a fi prea târziu 1
Partea I: Spiritul din aparat 15
1 Mintea şi maşinile: o foarte scurtă istorie a inteligenţei artificiale 17
2 Lumea insulară a triburilor inteligenţei artificiale 66
3 O mie de tăieturi cu hârtia: consecinţele neintenţionate ale inteligenţei artificiale 126
Partea a II-a: Viitorurile noastre 171
4 De la prezent la super-inteligenţa artificială: semnalele de avertizare 173
5 Prosperitate în a treia epocă a calculatoarelor: scenariul optimist 199
6 Cum să învăţăm să trăim cu milioane de tăieturi cu hârtia: scenariul pragmatic 229
7 Dinastia Rengong Zhineng: scenariul catastrofic 265
Partea a III-a: Rezolvarea problemelor 297
8 Pietricele şi bolovani: cum să reparăm viitorul inteligenţei artificiale 299
Mulţumiri 335
Bibliografie 339
Note 349

INTRODUCERE

ÎNAINTE DE A FI PREA TÂRZIU

Inteligenţa artificială face deja parte din viaţa noastră, însă nu sub forma pe care o aşteptasem cu toţii, magnifici

Inteligenţa artificială reprezintă structura discretă de rezistenţă a sistemelor noastre financiare, a reţelelor energetice şi a lanţurilor de aprovizionare comerciale. Ea reprezintă infrastructura invizibilă care ne conduce prin trafic, găseşte înţelesul corect în cuvintele noastre tastate greşit şi stabileşte ce ar trebui să cumpărăm, urmărim, ascultăm sau citim. Ea este tehnologia pe care se

construieşte viitorul nostru fiindcă se intersectează cu fiecare aspect din viaţa noastră: sănătatea şi medicina, locuinţele, agricultura, transporturile, sporturile şi chiar iubirea, sexul şi moartea.

Inteligenţa artificială nu e un trend în tehnologie, sau un cuvânt cheie, sau o distracţie de moment. Ea reprezintă a treia eră a informaticii. Ne aflăm în mijlocul unor transformări importante, asemănătoare cu cele pe care le-au trăit generaţiile din epoca revoluţiei industriale. La început, oamenii nu au putut să recunoască tranziţia prin care treceau pentru că schimbarea s-a petrecut treptat, dacă ne raportăm la durata lor de viaţă. La finalul procesului, lumea arăta diferit: Marea Britanie şi Statele Unite erau cele două puteri dominante în lume, având suficient capital industrial, militar şi politic pentru a influenţa parcursul următorului secol.

Toată lumea dezbate până la saturaţie inteligenţa artificială şi implicaţiile acesteia pentru viitorul nostru. Cunoaşteţi deja argumentele cele mai frecvente: roboţii ne vor lua locurile de muncă, roboţii vor da economia peste cap, roboţii vor sfârşi prin a ucide oameni. Dacă înlocuim „maşină” cu „robot” vom regăsi aceleaşi dezbateri pe care le aveau oamenii acum 200 de ani. E normal să ne gândim asupra impactului noilor tehnologii asupra slujbelor noastre şi a capacităţii noastre de a câştiga bani, deoarece vedem că multe industrii trec prin transformări profunde. E de înţeles de ce atunci când ne gândim la inteligenţa artificială, mintea noastră se gândeşte la exemple celebre precum HAL 9.000 din filmul *2000*: *Odiseea spaţială*, WOPR din *Simulări de război*, Skynet din *Terminator*, Roşie din *Familia Jetson*, Dolores din *Westworld* sau oricare dintre sutele de alte forme antropomorfizate ale inteligenţei artificiale din cultura pop. Dacă nu lucraţi direct în interiorul ecosistemului inteligenţei artificiale, viitorul vi se poate părea fie extraordinar, fie îngrozitor, dar asta din nişte motive complet greşite.

Persoanele neimplicate în cercetarea şi dezvoltării

inteligenţei artificiale nu pot să vadă în mod clar semnalele, motiv pentru care dezbaterea publică pe această temă se rezumă la roboţi care vor să domine lumea, sau reflectă un soi de optimism maniacal, îndrăzneţ. Lipsa nuanţelor constituie una dintre problemele inteligenţei artificiale încă de la începuturile acesteia: unii supraestimează dramatic domeniile în care ar putea fi aplicată, în vreme ce alţii susţin că va deveni o armă de neoprit.

Ştiu asta pentru că mi-am petrecut o bună parte din ultimul deceniu în domeniul cercetării inteligenţei artificiale şi m-am întâlnit cu oameni şi organizaţii atât din interiorul cât şi din afara ecosistemului inteligenţei artificiale. Am fost consultant pentru diferite companii de vârf din domeniu, printre care se numără Microsoft sau IBM. M-am întâlnit şi am consiliat diverse persoane cointeresate: investitori şi manageri de fonduri mutuale, lideri din Departamentul Apărării şi Departamentul de Stat, şi diferiţi parlamentari care cred că viitorul nu poate exista decât într-un cadru reglementat. Am avut şi sute de întâlniri cu cercetători şi ingineri care lucrează chiar în tranşee. Cei care lucrează direct cu inteligenţa artificială nu împărtăşesc decât rareori viziunile extreme – fie utopice, fie apocaliptice despre care auzim în presă.

Motivul e că cercetătorii din alte domenii ale ştiinţei, cei care chiar pun umărul la construirea viitorului inteligenţei artificiale vor să tempereze aşteptările. Depăşirea unor limite istorice se face cu răbdare, timp, bani şi statornicieuităm constant acest aspect. Aceşti oameni muncesc din greu, rezolvând pas cu pas probleme incredibil de complexe, şi reuşind uneori doar progrese foarte modeste. Ei sunt deştepţi, sociabili, şi pe baza experienţelor mele, nişte oameni înţelegători şi atenţi.

Aceştia muncesc în proporţie covârşitoare pentru nouă companii uriaşe din domeniul tehnologiei-Google, Amazon, Apple, IBM, Microsoft şi Facebook în Statele Unite, respectiv Baidu, Alibaba şi Tencent în China – care

construiesc inteligenţa artificială pentru a ne aduce tuturor un viitor mai bun, mai luminos. Cred cu tărie că liderii acestor companii sunt determinaţi de un simţ profund al altruismului şi de o dorinţă de a contribui la binele comunităţii: ei văd cu claritate potenţialul pe care-l oferă inteligenţa artificială pentru a îmbunătăţi sănătatea şi longevitatea, a rezolva problemele climatice şi a aduce bogăţia în casele a milioane de oameni. Vedem deja beneficiile reale şi pozitive ale muncii lor în toate domeniile şi în viaţa de zi cu zi.

Problema e că există şi forţe externe care acţionează asupra celor nouă giganţi tehnologici – şi, prin extensie, asupra celor care lucrează în interiorul ecosistemuluiobligându-i să-şi desconsidere bunele intenţii pentru viitorul tuturor. Există mulţi vinovaţi.

În Statele Unite, solicitările permanente ale pieţei şi aşteptările nerealiste în privinţa noilor produse şi servicii zădărnicesc orice planuri pe termen lung. Ne aşteptăm ca Google, Amazon, Apple, Facebook, Microsoft şi IBM să anunţe noi produse şi descoperiri legate de inteligenţa artificială la prezentările anuale, de parcă marile descoperiri ar putea fi planificate. Dacă aceste companii nu ne oferă produse mai strălucitoare decât cele de anul trecut, vorbim despre ele de parcă ar fi nişte rateuri. Sau punem la îndoială existenţa inteligenţei artificiale. Sau le punem în discuţie liderii. Nu am lăsat niciodată acestor companii un răgaz de câţiva ani în care să se pună pe treabă fără să le pretindem să ne uluiască la intervale regulate. Doamne-fereşte ca vreuna dintre aceste companii să se decidă să nu comunice nimic vreme de câteva luni- vom presupune că tăcerea lor ascunde un proiect care ne va lua prin surprindere.

Guvernul Statelor Unite nu are o strategie măreaţă în ce priveşte inteligenţa artificială sau viitorul nostru pe termen lung. Astfel, în locul unor strategii naţionale coordonate menite să sporească capacitatea organizaţională a guvernului, să ne construiască şi

întărească alianţele cu alte ţări, şi să ne pregătească armata pentru conflictele viitorului, Statele Unite au lăsat inteligenţa artificială la bunul plac al politicienilor care se perindă la guvernare. În loc să finanţeze un program de bază de cercetare în inteligenţa artificială, guvernul federal a externalizat efectiv cercetarea şi dezvoltarea sectorului privat şi capriciilor bursei de pe Wall Street. În loc să considere că inteligenţa artificială este o ocazie prin care se pot crea noi slujbe şi pune bazele creşterii economice, politicienii americani văd doar şomajul generalizat pe care l-ar putea provoca tehnologia. Ei dau la rândul lor vina pe giganţii tehnologiei din Statele Unite, în loc să invite aceste companii să ia parte la cele mai înalte niveluri ale planificării strategice ale guvernului (în măsura în care ea există). Pionierii inteligenţei artificiale nu au nicio alternativă în afara unei competiţii constante pentru a avea o legătură directă cu voi, cu mine, cu şcolile, spitalele, firmele şi oraşele noastre.

Suferim în Statele Unite de un miopism tragic. Trăim într-o paradigmă „prezentistă”, preocupându-ne următorii ani mult mai mult decât oricare alt interval. Gândirea prezentistă pune mare preţ pe realizările pe termen scurt ale tehnologiei, dar ne exonerează de răspunderea faţă de evoluţia tehnologiei şi rezultatele propriilor fapte. Uităm prea uşor că ceea ce facem în prezent ar putea avea consecinţe grave asupra viitorului. E aşadar de mirare că practic am externalizat viitorul inteligenţei artificiale către şase companii listate la bursă, ale căror realizări sunt spectaculoase, dar care au interese financiare care nu se suprapun întotdeauna cu un rezultat optim pentru libertăţile noastre, pentru comunităţile noastre şi pentru idealurile noastre democratice?

În China însă, evoluţia inteligenţei artificiale este strâns legată de marile ambiţii ale guvernului. China îşi construieşte cu rapiditate bazele pentru a deveni autoritatea mondială de necontestat în domeniu. În luna iulie 2017, guvernul chinez a prezentat planul de

dezvoltare a următoarei generaţii de inteligenţă artificială pentru a obţine până în anul 2030 supremaţia în domeniul inteligenţei artificiale, urmând ca cifra de afaceri a industriei naţionale să atingă cel puţin pragul de 150 de miliarde de dolari, 1 prin alocarea unei cote din fondul suveran de investiţii către noi laboratoare de cercetare şi companii nou înfiinţate, precum şi noi facultăţi menite să pregătească următoarea generaţie de specialişti chinezi în inteligenţă artificială.2 în luna octombrie a aceluiaşi an, preşedintele chinez Xi Jinping a explicat planurile sale pentru inteligenţa artificială şi analiza datelor într-un discurs amănunţit în faţa a mii de lideri ai partidului. Inteligenţa artificială, susţine Xi, va ajuta China să devină una dintre cele mai avansate economii din lume. Economia Chinei este deja de 30 de ori mai mare decât era în urmă cu doar trei decenii. Deşi Baidu, Tencent şi Alibaba sunt giganţi listaţi la bursă, ei se supun voinţei Beijingului, aidoma oricărei mari companii din China.

Populaţia uriaşă a Chinei, de 1, 4 miliarde de locuitori, îi dă acestei ţări controlul asupra celei mai mari, şi poate chiar celei mai importante resurse naturale din epoca inteligenţei artificiale: datele despre oameni. E nevoie de cantităţi considerabile de date pentru a îmbunătăţi algoritmii de recunoaştere a tiparelor – motiv pentru care sistemele chinezeşti de recunoaştere facială precum Megvii sau Sense Time sunt atât de atrăgătoare pentru investitori. Toate datele pe care cetăţenii chinezi le generează în timp ce vorbesc la telefon, fac cumpărături pe internet şi postează fotografii pe reţelele de socializare sunt folosite de Baidu, Alibaba şi Tencent pentru a crea sisteme performante de inteligenţă artificială. Un mare avantaj al Chinei este că nu are aceleaşi restricţii de securitate şi protecţie a datelor personale care ar putea limita progresul în acest domeniu în Statele Unite.

Trebuie să analizăm parcursul inteligenţei artificiale în contextul mai larg al marilor planuri de viitor ale Chinei. În luna aprilie 2018, Xi a ţinut un discurs important în

care-şi prezenta viziunea asupra Chinei ca super-putere mondială. Agenţia de presă de stat Xinhua a publicat fragmente din discurs, în care Xi amintea o nouă reţea de administrare a spaţiului cibernetic şi un internet care ar „transmite informaţii pozitive, ar respecta orientarea politică corectă şi ar îndruma opinia publică şi valorile în direcţia bună”. 3 Regulile autoritare pe care China încearcă să le impună în restul lumii constituie o abatere de la libertatea de exprimare, economia de piaţă şi împărţirea puterii, valori pe care le îndrăgim în lumea occidentală.

Inteligenţa artificială este doar o componentă a unor legi care urmăresc să controleze toate informaţiile generate pe teritoriul Chinei şi să monitorizeze datele propriilor cetăţeni, precum şi ale cetăţenilor diferiţilor săi parteneri strategici. Una dintre aceste legi solicită tuturor companiilor din străinătate să stocheze datele referitoare la cetăţenii chinezi în servere aflate pe teritoriul chinezesc. Din cauza acestei reglementări, agenţiile de siguranţă ale guvernului pot accesa după bunul plac aceste date. Un alt proiect – Cloud-ul poliţiei chinezea fost creat pentru a monitoriza şi urmări persoanele cu probleme psihice, disidenţii declaraţi şi o minoritate etnică musulmanăuigurii. În luna august 2018, Naţiunile Unite au anunţat că există surse credibile conform cărora China deţine milioane de uiguri în lagăre secrete în regiunile vestice îndepărtate ale Chinei.4 Programul integrat de monitorizare al Chinei foloseşte inteligenţa artificială pentru a depista eventuale abateri – detectând, de exemplu, dacă cineva îşi plăteşte facturile cu întârziere. Sistemul de credite sociale bazat pe inteligenţa artificială a fost creat, conform unui slogan din documentele oficiale de planificare, pentru a instaura o societate lipsită de probleme „permiţând cetăţenilor care respectă legea să prospere, împiedicându-i însă pe cei discreditaţi să facă chiar şi un pas”. 5 Pentru a putea promova „încrederea”, cetăţenii sunt evaluaţi în funcţie de diferite criterii cum ar

fi faptele de curaj (puncte în plus) sau amenzile de circulaţie (puncte în minus). Cei cu scoruri mai mici îşi găsesc mai greu o slujbă, o casă sau un loc în şcoală pentru copii. În unele oraşe, fotografiile cetăţenilor ascultători sunt afişate în public.6 în alte oraşe, de pildă în Shandong, cetăţenii care trec strada prin locuri nepermise sunt afişaţi în public pe panouri digitale şi pe Weibo, o reţea populară de socializare.7 Dacă vi se pare neverosimil, aduceţi-vă aminte că în China s-a putut decreta ca toate familiile să nu facă decât un singur copil pentru a-şi reduce forţat populaţia.

Aceste politici şi iniţiative provin din anturajul apropiat preşedintelui Xi Jinping, format din persoane care în ultimul deceniu au acordat o atenţie extraordinară repoziţionării şi reconstruirii Chinei ca hegemon mondial. China este mai autoritară în zilele noastre decât sub orice alt lider de după Mao Zedung, iar progresul în domeniul inteligenţei artificiale şi folosirea acesteia constituie elemente esenţiale ale supravieţuirii regimului. Iniţiativa Noului Drum al Mătăsii constituie o strategie geoeconomică deghizată sub forma unui plan vast de dezvoltare a infrastructurii de-a lungul fostului Drum al Mătăsii care unea odinioară China cu Europa prin Orientul Mijlociu şi Africa. China nu construieşte doar poduri şi autostrăzi-China exportă şi tehnologie de supraveghere şi strânge date, sporind influenţa Partidului Comunist Chinez în lumea, în faţa ordinii liberale actuale. Interconectarea energetică globală este o altă iniţiativă naţională a lui Xi care urmăreşte crearea primei reţele globale de electricitate, sub umbrela Chinei. China dispune deja de o tehnologie revoluţionară cu tensiuni foarte înalte, care-i permite să transporte curent electric din îndepărtatele regiuni vestice către Shanghai, şi încheie diverse acorduri comerciale pentru a furniza energie electrică ţărilor vecine.

Aceste planuri de viitor, dar şi multe altele, constituie modalităţi inteligente de a obţine o putere subtilă pe o

perioadă lungă de timp. Ele constituie o mutare strategică de geniu a lui Xi, al cărui Partid Comunist Chinez a votat în luna martie 2018 să elimine limitele de mandat, permiţându-i acestuia să rămână preşedinte pe viaţă. Scopul lui Xi e cât se poate de evident: să făurească o nouă ordine mondială în care China să devină liderul de facto. Cu toate acestea, într-o perioadă de expansiune a influenţei diplomatice chineze, Statele Unite şi-au întors spatele alianţelor globale de durată, iar preşedintele Trump a înălţat o cortină nouă, de bambus.

Viitorul inteligenţei artificiale urmează două parcursuri de dezvoltare aflate adesea în conflict cu interesele omenirii. Demersurile chinezeşti în domeniul inteligenţei artificiale sunt parte a unei tentative coordonate de a instaura o nouă ordine mondială sub bagheta preşedintelui Xi, în vreme ce în America principalii factori determinanţi sunt forţele pieţei şi consumismul. Această dihotomie reprezintă un punct nevralgic pentru noi toţi. Rezolvarea ei reprezintă partea esenţială a problemelor care ne pândesc în viitor, precum şi motivul pentru care am scris această carte. Cei Nouă Titani ar putea urmări aceleaşi obiective nobile, comune – decodarea secretelor care ar permite construirea unor sisteme capabile să producă o gândire asemănătoare celei omeneşti – rezultatul final al acestei activităţi ar putea afecta însă iremediabil omenirea.

Eu cred că inteligenţa artificială este, la bază, o forţă a binelui care va contribui la un trai mai bun pentru generaţiile viitoare şi ne va ajuta să atingem visurile despre viitor.

Sunt şi o fire pragmatică, însă. Cu toţii ştim că până şi oamenii cu intenţiile cele mai bune pot să facă mult rău fără să vrea. În domeniul tehnologiei, şi mai ales când vine vorba de inteligenţa artificială, trebuie să avem constant în vedere că aceste tehnologii pot fi folosite atât aşa cum au fost proiectate, cât şi abuzate în moduri neprevăzute. În prezent, dar şi în viitorul pe care-l putem întrezări, această

constatare este deosebit de importantă deoarece inteligenţa artificială afectează toate domeniile: economia globală, forţa de muncă, agricultura, transporturile, sistemul bancar, monitorizarea mediului, educaţia, armata şi siguranţa naţională. Acestea sunt motivele pentru care, dacă inteligenţa artificială îşi va păstra parcursul actual în Statele Unite şi China, anul 2069 ar putea fi radical diferit faţă de cum îl anticipăm în 2019. Pe măsură ce structurile şi sistemele care guvernează societatea ajung să depindă de inteligenţa artificială, vom constata că pentru maşini deciziile luate în locul nostru sunt complet rezonabile doar că nouă ele nu ni se vor părea astfel.

Ca oameni, ne pierdem rapid înţelegerea asupra fenomenului, chiar în timp ce maşinile de abia se trezesc. Am depăşit câteva pietre de hotar în evoluţia tehnică şi geopolitică a inteligenţei artificiale, însă cu fiecare nou progres inteligenţa artificială devine din ce în ce mai opacă pentru publicul larg. Modul în care datele noastre sunt extrase şi prelucrate devine mai puţin clar, în timp ce capacitatea noastră de a înţelege felul în care sistemele autonome iau decizii devine din ce în ce mai redusă. Am ajuns astfel la o prăpastie în ce priveşte înţelegerea felului în care inteligenţa artificială ne afectează viaţa cotidiană, iar această prăpastie se măreşte exponenţial pe măsură ce avansăm cu anii şi deceniile în viitor. Misiunea acestei cărţi este să aducă în atenţia publicului dezbaterea privind inteligenţa artificială şi să vă pregătească pentru ce urmează – şi să vă fac să înţelegeţi care vor fi consecinţele în viaţa de zi cu zi pentru fiecare dintre noi, înainte de a fi prea târziu.

Omenirea se confruntă la propriu cu o criză existenţială, pentru că nimeni nu răspunde la o întrebare simplă care s-a aflat la baza inteligenţei artificiale încă de la început: ce se petrece cu o societate atunci când transferăm puterea unui sistem construit de un grup restrâns de oameni, care e menit să ia decizii în numele tuturor? Ce se întâmplă când aceste decizii favorizează

piaţa, sau un partid politic ambiţios? Răspunsul se reflectă în oportunităţile viitoare pe care le avem, în felul în care ni se refuză accesul, în convenţiile sociale din cadrul societăţilor noastre, în regulile pe baza cărora operează economiile noastre şi chiar în felul în care ne raportăm la ceilalţi.

Aceasta nu este o carte despre dezbaterile obişnuite din domeniu. Ea e atât un semnal de alarmă, cât şi un tipar pentru un viitor mai bun. Cartea pune la îndoială aversiunea faţă de planurile pe termen lung pe care o simţim în Statele Unite şi subliniază lipsa de pregătire în domeniu pe care o vădesc firmele, şcolile şi guvernul nostru. Cartea oferă o imagine sinceră a strategiilor interconectate geopolitice, economice şi diplomatice ale Chinei, care-şi urmăreşte propria sa viziune asupra unei noi ordini mondiale. Cartea e şi un îndemn către o conducere eroică în împrejurări deosebit de grele. Veţi afla din paginile acestei cărţi că viitorul nostru are nevoie de un erou.

Urmează un apel scris în trei părţi. Din prima parte veţi afla ce este inteligenţa artificială şi rolul pe care-l joacă cei Nouă Titani în evoluţia acesteia. Vom analiza îndeaproape şi particularităţile celor nouă companii în America, respectiv ale Baidu, Alibaba şi Tencent în China. În partea a doua veţi descoperi trei scenarii amănunţite, plauzibile despre viitorii 50 de ani, pe măsură ce inteligenţa artificială progresează. Cele trei scenarii pe care le veţi citi oscilează între o abordare optimistă, pragmatică şi catastrofică, ele revelându-ne atât riscurile cât şi oportunităţile pe care le incumbă trecerea de la inteligenţa artificială îngustă la inteligenţa artificială generală şi apoi la superinteligenţa artificială. Aceste scenarii sunt intenseele sunt rezultatul unor modele analitice, permiţându-vă să aveţi o perspectivă viscerală asupra evoluţiei potenţiale a inteligenţei artificiale şi a schimbărilor care le vor suferi ca urmare vieţile noastre, în partea a treia voi prezenta soluţii tactice şi strategice la

toate problemele pe care le-am identificat în scenarii, dar şi un plan concret de a reporni prezentul. Partea a treia este menită să ne redea imboldul de a acţiona, ea conţinând recomandări concrete pentru guvernele noastre, liderii celor nouă titani, şi chiar pentru voi.

Orice persoană aflată în viaţă în ziua de azi poate juca un rol esenţial în viitorul inteligenţei artificiale. Deciziile pe care le luăm acum în privinţa inteligenţei artificialechiar şi cele aparent banalevor modifica pentru totdeauna parcursul istoriei omenirii. Pe măsură ce maşinile se trezesc, ne-am putea da seama că, în ciuda speranţelor şi ambiţiilor noastre altruiste, sistemele inteligenţei artificiale ar putea reprezenta o catastrofă pentru omenire.

Nu e însă obligatoriu ca lucrurile să se petreacă astfel.

Cei Nouă Titani nu sunt personajele negative ale acestei poveşti. De fapt, aceste companii reprezintă cea mai bună speranţă a noastră pentru viitor.

Daţi pagina. Nu putem să stăm şi să aşteptăm viitorul. Inteligenţa artificială e deja printre noi.

PARTEA I

Spiritul din aparat

CAPITOLUL UNU

MINTEA ŞI MAŞINILE: O FOARTE SCURTĂ ISTORIE A INTELIGENTEI ARTIFICIALE

7

Originile inteligenţei artificiale moderne se întind în urmă cu

X.X. sute de ani, cu mult timp înainte ca cei Nouă Titani să construiască agenţi ai inteligenţei artificiale cu nume ca Siri, Alexa sau echivalentul lor chinezesc, Tiân Mão. În tot acest timp, inteligenţa artificială nu a avut o singură definiţie, aşa cum au alte tehnologii. Descrierea concretă a inteligenţei artificiale nu e atât de uşoară, şi asta pentru că ea reprezintă multe lucruri diferite, chiar şi în timp ce domeniul continuă să crească. Ceea ce se

considera în anii ’50 inteligenţă artificialăun calculator capabil să efectueze împărţirinu ni se mai pare vreo tehnologie avansată în zilele noastre. Acest fenomen este cunoscut sub numele de „paradoxul ciudat” de îndată ce noile tehnologii sunt inventate şi devin parte a vieţii cotidiene, ele devin invizibile pentru noi. Nu ni le mai reprezentăm ca fiind parte a inteligenţei artificiale.

În forma sa cea mai elementară, inteligenţa artificială reprezintă un sistem capabil să ia decizii autonome. Activităţile desfăşurate de inteligenţa artificială copiază sau imită inteligenţa omenească, sub forma recunoaşterii sunetelor, obiectelor, rezolvării problemelor, înţelegerii limbajului şi a folosirii strategiei pentru atingerea unor obiective. Unele sisteme sunt gigantice, putând efectua rapid milioane de calcule, în vreme ce altele sunt dedicate unei singure misiuni, cum ar fi depistarea cuvintelor obscene din e-mailuri.

Am revenit întotdeauna la acelaşi set de întrebări: sunt maşinile în stare să gândească? Ce ar însemna ca o maşină să „gândească”? Ce înseamnă pentru noi să gândim? Ce este gândirea? Cum am putea şticu siguranţă, şi fără îndoială – că avem gânduri originale? Aceste întrebări ne-au însoţit pe parcursul secolelor şi ele sunt esenţiale atât în istoria cât şi în viitorul inteligenţei artificiale.

Problema studierii felului în care atât maşinile cât şi oamenii gândesc e faptul că cuvântul „a gândi” este strâns legat de noţiunea de „minte”. Dicţionarul Merriam-Webster defineşte „a gândi” ca fiind „a forma sau a avea în minte” în vreme ce dicţionarul Oxford explică că înseamnă „a folosi propria minte în mod activ pentru a înlănţui idei”. Dacă căutăm cuvântul „minte”, atât Merriam-Webster cât şi Oxford oferă o definiţie în contextul „conştienţei”. Dar ce este conştienţa? Conform ambelor dicţionare, ea reprezintă starea de a fi conştient şi de a reacţiona. Diferite grupuri – psihologi, neurobiologi, filosofi, teologi, eticieni şi informaticieni – abordări separate când vine

vorba de noţiunea de gândire.

Când folosiţi Alexa pentru a găsi o masă la restaurantul preferat, atât voi cât şi ea sunteţi conştienţi şi reactivi când vine vorba de a mânca, chiar dacă Alexa nu a simţit niciodată senzaţia pe care o provoacă dinţilor textura unui măr crocant sau înţepăturile efervescente ale apei minerale pe limbă sau cum se lipeşte untul de arahide de cerul gurii. Întrebaţi-o pe Alexa să vă descrie calităţile acestor mâncăruri, şi ea vă va oferi detalii care oglindesc propriile voastre experienţe. Alexa nu are gurăşi atunci cum ar putea ea să perceapă la fel ca voi mâncarea?

Sunteţi o persoană unică din punct de vedere biologic, ale cărei glande salivare şi papile gustative nu sunt poziţionate la fel cu ale mele. Cu toate acestea, amândoi am învăţat ce este un măr, caracteristicile generale ale gustului de măr, textura lui şi mirosul lui. Pe parcursul vieţii, am învăţat să recunoaştem ce este un măr prin consolidarea informaţieicineva ne-a învăţat cum arată un măr, la ce serveşte, şi prin ce se deosebeşte de alte fructe. După aceea, cu timpul şi fără să ne dăm seama conştient de asta, sistemele noastre biologice autonome de recunoaştere a tiparelor au început să poată să stabilească cu mare precizie dacă un obiect este un măr, chiar dacă avem la dispoziţie doar câteva dintre elementele necesare. Dacă vedeţi un contur alb-negru, bidimensional al unui măr, veţi şti ce este – chiar dacă nu are miros, gust, textură şi toate celelalte date pe baza cărora creierul vostru poate identifica un măr. Felul în care voi şi Alexa aţi învăţat despre mere e mai asemănător decât credeţi.

Alexa e competentă, dar e ea *inteligentă*? E nevoie ca percepţia ei maşinală să bifeze toate calităţile percepţiei omeneşti pentru ca noi să putem accepta că felul ei de „gândire” îl oglindeşte fidel pe al nostru? Dr. Benjamin Bloom, cercetător în psihologia educaţională, şi-a petrecut cea mai mare parte a carierei academice studiind şi clasificând stările de gândire. El a publicat în anul 1956 taxonomia lui Bloom, care definea obiectivele de învăţare

şi nivelurile realizărilor obţinute în educaţie. Stratul de bază îl reprezintă reţinerea datelor şi conceptelor elementare, urmat apoi de înţelegerea ideilor; aplicarea cunoştinţelor în situaţii noi; analizarea informaţiilor prin experimentare şi conexiuni între idei; evaluarea, susţinerea şi criticarea informaţiei; totul culminând cu crearea unor opere originale. Din copilărie, ne concentrăm mai întâi pe înmagazinare şi înţelegere. Spre exemplu, trebuie mai întâi să învăţăm că o sticlă conţine lapte înainte de a înţelege că sticla respectivă are o parte din faţă şi una din spate, chiar dacă nu o putem vedea.

Această ierarhie se regăseşte şi în felul în care calculatoarele învaţă. În 2017, un sistem de inteligenţă artificială denumit Amper a compus şi produs muzică originală pentru un album intitulat I AM AI. Instrumentaţia şi percuţia au fost generate de Amper, folosind parametrii iniţiali precum genul muzical, dispoziţia sau lungimea melodiei, pentru a crea un cântec în doar câteva minute. Un artist viu, Taryn Southern, a colaborat cu Amper la crearea albumului, iar rezultatul a fost o baladă emoţionantă denumită „Break Free” care a strâns peste 1, 6 milioane de vizualizări pe You-Tube şi a devenit un hit la radio. Înainte ca Amper să poată crea acest cântec, a trebuit mai întâi să înveţe care sunt caracteristicile unei balade, precum şi date concrete, referitoare la lungimile notelor şi ritmurilor şi cum să recunoască mii de tipare în muzică (acorduri, accente ritmice, secvenţe armonice).

Genul de creativitate demonstrată de Amper reprezintă vârful taxonomiei lui Bloom, dar avem aici de-a face doar cu un proces învăţat mecanic? E un exemplu de creativitate omenească? Sau e o creativitate în cu totul alt sens? se gândea Amper la muzică, aşa cum o poate face un compozitor viu? S-ar putea susţine că „creierul” lui Amper – o reţea neurală formată din algoritmi şi date într-un recipient – nu e chiar atât de diferit de creierul lui Beethoven, constituit din neuroni organici care folosesc date şi recunosc tipare într-un recipient, adică la el în cap.

A trecut Amper oare printr-un proces creativ complet diferit de cel al lui Beethoven atunci când a compus Simfonia a 5-a, cea care începe celebru prin ta-ta-ta-TAM, ta-ta-ta-TAM înainte de a trece de la o gamă majoră la una minoră? Aceste prime patru note sunt urmate de o secvenţă armonică, de game, arpegii şi alte ingrediente brute ale oricărei compoziţii. Dacă ascultaţi cu atenţie *scherzoul*, înainte de final, veţi vedea un împrumut evident din simfonia nr. 40 a lui Mozart, scrisă cu 20 de ani în urmă, în 1788. Mozart a fost influenţat de rivalul lui, Antonio Salieri, şi de prietenul său, Franz Joseph Haydn, care au fost la rândul lor influenţaţi de operele unor compozitori mai vechi precum Johann Sebastián Bach, Antonio Vivaldi sau Henry Purceii, care au scris muzică de la mijlocul secolului XVII până la mijlocul secolului XVIII. în muzica lor se pot detecta pasaje din compozitori şi mai vechi, de la 1400 – 1600, cum ar fi Jacques Arcadelt, Jean Mouton şi Johannes Ockeghem. *Şi ei* au fost influenţaţi de primii compozitori medievali – şi am putea continua cu influenţa până la prima compoziţie intitulată epitaful Seikilos, gravat în secolul I e.n. Într-o coloană de marmură într-un mormânt din Turcia zilelor noastre. Am putea continua chiar şi mai mult în timp, până la primele flaute primitive, sculptate din os şi fildeş acum 43.000 de ani. Chiar şi înaintea acestei perioade, cercetătorii cred că strămoşii noştri au cântat înainte de a vorbi.8

Felul în care gândim este rezultatul a milioane de ani de evoluţie. Felul în care funcţionează inteligenţa artificială modernă se bazează în mod similar pe un lung parcurs al evoluţiei, până la primii matematicieni, filosofi şi oameni de ştiinţă ai antichităţii. Deşi ar putea părea că omenirea şi maşinile au avut parcursuri diferite, evoluţia noastră a fost întotdeauna strâns legată de a lor. *Homo sapiens* a învăţat să-şi cunoască mediul, a transmis anumite trăsături generaţiilor viitoare, s-a înmulţit şi diversificat din cauză că a inventat tehnologii avansate, cum ar fi agricultura, uneltele de vânătoare sau penicilina.

A fost nevoie de 11.000 de ani pentru ca de la cei 6 milioane de locuitori ai planetei din neolitic să se ajungă la populaţia de 7 miliarde din zilele noastre.9 Ecosistemul în care evoluează sistemele inteligenţei artificiale – învăţarea, datele, algoritmii, procesoarele, aparatele şi reţelele neurale – se îmbunătăţeşte exponenţial. Nu e decât o chestiune de decenii până când sistemele inteligenţei artificiale se vor răspândi şi integra în toate aspectele vieţilor noastre.

Dacă Alexa poate percepe un măr la fel ca noi, şi dacă muzica lui Amper chiar e „originală” – acestea sunt întrebări care vizează mai mult felul în care definim gândirea. Inteligenţa artificială din ziua de azi reprezintă un amalgam la care au contribuit de-a lungul mileniilor filosofi, matematicieni, oameni de ştiinţă, roboticieni, artişti şi teologi. Misiunea lor – dar şi a noastră, în acest capitol – este să înţeleagă legătura dintre gândire şi recipientele acesteia. Care este legătura dintre mintea omenească şi maşinile construite de cei Nouă Titani în China şi Statele Unite-dacă *nu cumva aceste maşini sunt construite în ciuda minţii omeneşti7*.

Există o minte în interiorul maşinii?

Primele baze ale inteligenţei artificiale provin din Grecia antică, odată cu originile filosofiei, logicii şi matematicii. În multe din scrierile lui Platon, Socrate spune „cunoaşte-te pe tine însuţi”, iar el susţinea că pentru a putea să deveniţi mai buni şi a lua decizii corecte, trebuie mai întâi să vă cunoaşteţi caracterul. Printre alte contribuţii, Aristotel a inventat logica silogistică şi primul sistem formal de raţionament deductiv. Cam în aceeaşi perioadă, matematicianul Euclid a inventat o cale pentru a găsi cel mai mare numitor comun între două numere, creând astfel primul algoritm. Din opera lor s-au născut două idei importante: că anumite sisteme fizice pot opera pe baza unui set de reguli logice, şi că însăşi gândirea omenească ar putea fi un sistem simbolic. De aici au pornit sute de ani de investigaţii ale filosofilor, teologilor şi

oamenilor de ştiinţă. E corpul un aparat complex? Un întreg compus din sute de alte sisteme care funcţionează laolaltă, ca o pendulă? Dar mintea? E şi ea o maşinărie complicată? Sau poate ceva cu totul diferit? Nu aveam cum să dovedim sau să infirmăm existenţa unui algoritm divin, sau o legătură între minte şi lumea fizică.

În anul 1560, un ceasornicar pe nume Juanelo Turnano a creat un mic călugăr drept ofrandă adusă bisericii, în numele regelui Filip al II-lea al Spaniei, al cărui fiu îşi revenise în mod spectaculos după un traumatism la cap.10 Acest călugăr avea puteri uluitoare – putea să meargă pe masă, să ridice un crucifix şi un rozariu, să se bată în piept în semn de penitenţă şi să-şi mişte buzele ca şi cum s-ar ruga. El a fost primul *automaton* – o reprezentare mecanică a unei fiinţe vii. Deşi cuvântul „robot” nu exista încă, călugărul reprezenta o invenţie remarcabilă, care trebuie că a şocat şi încurcat publicul din vremea aceea. Nimeni nu şi-ar fi putut probabil închipui că un mic automat ar putea într-un viitor îndepărtat nu doar să imite mişcările cele mai elementare ci ar putea să înlocuiască oamenii în fabrici, laboratoare sau conversaţii de bucătărie.

Micuţul călugăr a inspirat prima generaţie de roboticieni, al căror obiectiv a fost să creeze maşinării din ce în ce mai complexe care să imite oamenii: automatele au devenit curând capabile să scrie, danseze sau picteze. Aici a intervenit un grup de filosofi care au început să se întrebe ce anume înseamnă a fi om. Dacă este posibil să construim automate care imită comportamentul omenesc, atunci sunt oamenii automate de fabricaţie divină? Sau suntem sisteme complexe capabile de raţiune şi gândire proprie?

Filosoful politic englez Thomas Hobbes a descris raţiunea omenească drept o formă de calcul în opera sa de *Corpore*, parte a măreţei sale trilogii despre ştiinţele naturale, psihologie şi politică. Acesta scria în 1655: „Prin raţiune, înţeleg calcul. Iar a calcula înseamnă a aduna

multe lucruri împreună în acelaşi timp, sau a şti ce rămâne când scădem un lucru dintr-un altul. A raţiona este prin urmare acelaşi lucru cu a aduna sau a scădea”. 11 Dar cum am putea şti dacă avem şi liber arbitru în tot acest timp?

În timp ce Hobbes îşi scria prima parte a trilogiei, filosoful francez René Descartes publica opera *Meditaţii metafizice*, întrebându-se dacă putem şti cu siguranţă că ceea ce percepem este adevărat. Cum putem să ne verificăm conştienţa? De ce dovezi am avea nevoie pentru a concluziona că gândurile ne aparţin, şi că lumea din jurul nostru e reală? Descartes era un raţionalist, crezând că putem obţine faptele prin deducere. El a pus bazele unui experiment celebru de gândire. El le-a cerut cititorilor să-şi imagineze un demon care creează intenţionat o iluzie a lumii lor. Dacă senzaţia fizică, senzorială a cititorului de a înota într-un lac nu ar fi nimic altceva decât un construct al demonului, atunci acesta nu ar putea să *ştie* cu adevărat că înoată. În perspectiva lui Descartes, însă, dacă cititorul este înzestrat cu conştientizarea propriei existenţe, atunci acesta îndeplineşte criteriile pentru a fi o fiinţă care cunoaşte. „Sunt, exist, ori de câte ori rostesc asta, sau mintea mea concepe, atunci trebuie că e adevărat”, scria el.12 Cu alte cuvinte, faptul că existăm este dincolo de orice îndoială, chiar dacă în mijlocul nostru s-ar afla un demon care să vrea să ne amăgească. Sau, *gândesc*, *deci exist*.

Mai târziu, în opera de *Homine*, Descartes susţinea că oamenii ar putea construi probabil un automat – în exemplul lui, un mic animal – care să nu poată fi distins de original. Chiar dacă am putea cândva crea un om mecanic, acesta nu ar putea fi credibil, susţinea Descartes, pentru că nu ar avea minte, prin urmare nici suflet. Spre deosebire de oameni, un aparat nu ar putea îndeplini criteriul cunoaşterii – nu ar putea fi niciodată conştient de sine aşa cum suntem noi. Pentru Descartes, conştiinţa se manifestă intern – sufletul era fantoma din aparatul

complicat al corpului nostru.13

Cu câteva decenii mai târziu, matematicianul şi filosoful Gottfried Wilhelm von Leibniz a studiat noţiunea conform căreia sufletul omenesc ar fi la rândul său programat, susţinând că mintea însăşi reprezintă doar un recipient. Dumnezeu ar fi creat sufletul şi corpul în mod armonios. Corpul ar putea fi o maşinărie complexă, însă ea este programată conform unor instrucţiuni divine. Mâinile noastre se mişcă atunci când vrem noi, dar nu noi am creat sau inventat toate mecanismele care permit această mişcare. Dacă simţim durerea sau plăcerea, aceste senzaţii sunt rezultatul unui sistem preprogramat, al unei linii neîntrerupte de comunicare între minte şi corp.

Leibniz a venit cu propriul lui exerciţiu de gândire pentru a demonstra că gândirea şi percepţia sunt legate inextricabil de esenţa omenească. Imaginaţi-vă că intraţi într-o moară. Clădirea este un recipient care adăposteşte aparate, materii prime şi muncitori. Este un sistem un sistem complex cu părţi care operează armonios pentru atingerea unui singur scop, dar nu ar putea fi niciodată o minte. „Nu găsim acolo decât roţi dinţate şi mânere care se activează unul pe celălalt, dar nimic care să poată explica percepţia” scria Leibniz. „Prin urmare, percepţia trebuie căutată în substanţele simple, şi niciodată în lucrurile complexe cum ar fi aparatele”. Argumentul lui era că indiferent cât de avansată ar fi fost moara, aparatul sau automatele, oamenii nu vor putea niciodată să construiască o maşină capabilă să gândească sau să perceapă.14

Cu toate acestea, Leibniz era fascinat de noţiunea reproducerii anumitor aspecte ale gândirii. Cu câteva decenii în urmă, un scriitor englez obscur pe nume Richard Braithwait, care a scris câteva cărţi despre comportarea în societate a adus aminte în treacăt „calculatoarele”, referindu-se la oamenii care se pricep să calculeze rapid şi precis.15 între timp, matematicianul şi inventatorul francez Blaise Pascal, care a pus bazele

probabilităţii aşa cum o cunoaştem în zilele noastre, se interesa de automatizarea calculelor. Pascal îşi urmărise tatăl care îşi calcula meticulos impozitele şi voia să-i uşureze procesul. Prin urmare, Pascal a început să lucreze la un calculator automat, cu roţi mecanice şi cadrane mobile.16 Calculatorul funcţiona, el inspirându-l pe Leibniz să-şi corecteze gândirea: maşinile nu vor avea niciodată un suflet, cu toate acestea, ar putea deveni cândva posibil să construim o maşină capabilă de gândire logică la nivelul omului. În 1673, Leibniz a descris un „Calculator bazat pe paşi”, un nou tip de maşină de calcul care lua decizii în baza unui sistem binar.17 Maşina era un soi de masă de biliard, cu mingi, găuri şi canale, iar maşina deschidea găurile folosind o serie de 1 (deschis) şi o (închis).

Calculatorul teoretic bazat pe paşi al lui Leibniz a pus bazele unor noi teorii, printre care se numără şi noţiunea că dacă gândirea logică ar putea fi redusă la simboluri şi, ca urmare, ar putea fi analizată ca un sistem de calcul, şi dacă problemele geometrice ar putea fi rezolvate folosind simboluri şi numere, atunci totul ar putea fi redus la biţi – inclusiv comportamentul omenesc. Aceasta reprezenta o abatere semnificativă faţă de filosofii din vechime: maşinăriile viitorului ar putea reproduce gândirea omenească fără să încalce providenţa divină. Gândirea nu presupunea în mod obligatoriu percepţie, simţuri sau un suflet. Leibniz îşi imaginase un calculator capabil să rezolve probleme generale, chiar şi cele care nu sunt de ordin matematic. Şi el a presupus că limbajul se poate reduce la nivelul conceptelor elementare ale matematicii şi ştiinţei, ca parte a unui limbaj universal.18

Urmează pur şi simplu mintea şi aparatul un algoritm?

Dacă Leibniz avea dreptate – iar oamenii sunt nişte aparate cu suflete care vor inventa cândva maşini lipsite de suflet capabile de o gândire sofisticată, inimaginabilă – atunci ar putea exista o clasă binară de aparate pe pământ: noi şi ele. Dezbaterea era însă de abia la început.

În 1738, Jacques de Vaucanson, artist şi inventator, a construit pentru Academia Franceză de Ştiinţe o serie de automate, printre care se număra şi o raţă complexă, verosimilă. Ea nu doar că imita mişcările unei raţe vii, dând din aripi şi mâncând boabe – ci putea imita chiar şi digestia. Filosofii au avut ce diseca: dacă arăta ca o raţă, măcănea ca o raţă, era oare o raţă? Dacă credem că raţa are un altfel de suflet, ar fi suficient pentru a putea dovedi că raţa avea conştiinţă de sine, cu tot ceea ce ar presupune asta?

Filosoful scoţian David Hume a respins ideea că recunoaşterea existenţei constituie în sine dovada conştienţei. Spre deosebire de Descartes, Hume era un empirist. El a dezvoltat un nou cadru de cercetare bazat pe fapte observabile şi argumente logice. În timp ce de Vaucanson îşi prezenta raţa care digeră – şi cu mult timp ca vreun om să fi adus în discuţie inteligenţa artificială – Hume scria, în *Un tratat despre natura umană*: „Raţiunea este, şi ar trebui să fie, roaba pasiunilor”. În acest caz, Hume înţelegea prin „pasiuni” „motivaţii iraţionale” şi faptul că stimulii se află la baza comportamentului nostru, nu logica abstractă. Dacă impresiile sunt doar felul în care percepem lucrurile pe care le putem vedea, atinge, simţi, gusta şi mirosi, iar ideile sunt percepţii ale lucrurilor cu care nu intrăm direct în contact, Hume era de părere că existenţa şi înţelegerea lumii din jurul nostru se bazează pe un construct al percepţiei umane.

Odată cu noile progrese ale automatelor, care deveneau din ce în ce mai realiste, şi pe măsură ce ideea calculatoarelor ca maşini de gândit era luată din ce în ce mai serios, doctorul şi filosoful francez Julien Offray de La Mettrie a efectuat un studiu radical – şi scandalos – al oamenilor, animalelor şi automatelor, într-un articol publicat anonim în 1747, acesta concluziona că oamenii sunt foarte asemănători cu animalele, şi că dacă un hominin ar fi „pregătit cum trebuie”, acesta ar putea stăpâni limbajul oamenilor. La Mettrie e de părere şi că

oamenii şi animalele sunt doar nişte aparate, mânate de instinct şi experienţă. „Corpul omenesc e o maşină care-şi întoarce singură arcurile;…sufletul e doar un principiu al mişcării, sau o parte sensibilă şi materială a creierului”. 19

Ideea că oamenii ar fi doar nişte aparate bazate pe materie – nişte rotiţe care îndeplinesc un set de funcţii – presupune că nu suntem unici, sau speciali. Ea mai presupune şi faptul că am putea fi programabili. Dacă ar fi adevărat, şi dacă până în acest moment am putut crea raţe şi călugări care să arate ca originalul, atunci ar trebui să rezulte că într-o bună zi, oamenii s-ar putea reproduce pe ei înşişi – şi să construiască diverse aparate inteligente, capabile de gândire.

Putem construi o maşină care să gândească?

În anii 1830, matematicienii, inginerii şi oamenii de ştiinţă au început să meşterească, sperând că vor putea crea maşini capabile să facă aceleaşi calcule ca şi „calculatoarele” umane. Matematicienul Ada Lovelace şi omul de ştiinţă Charles Babbage au inventat o maşină denumită „Motorul diferenţial” şi apoi au imaginat un „Motor analitic” care folosea o serie de etape prestabilite pentru a rezolva probleme matematice. Lui Babbage nu i se părea că aparatul ar putea fi capabil de altceva decât de a calcula numerele. Lovelace a fost cea care, în notele de subsol ale unui articol ştiinţific pe care îl traducea, s-a abătut în mod genial speculând că o variantă mai puternică a Motorului ar putea fi folosită în alte moduri.20 Dacă maşina ar putea manipula simboluri, iar acestea ar putea la rândul lor să reprezinte diferite lucruri (spre exemplu note muzicale) atunci Motorul ar putea fi folosit pentru a „gândi” în afara domeniului matematicii. Deşi ea nu credea că un calculator ar putea vreodată să devină capabil să gândească pe cont propriu, ea şi-a imaginat un sistem complex care putea urma nişte instrucţiuni şi să imite o bună parte din activităţile zilnice ale oamenilor. Unora li se părea complet banal la vremea respectivă, însă Ada reuşise să scrie primul program complet pentru un

calculator puternic al viitoruluicu decenii înainte ca măcar becul să fi fost inventat.

O sută de mile mai la nord de universitatea Cambridge unde lucraseră Lovelace şi Babbage, un tânăr matematician auto-didact pe nume George Boole se plimba pe un câmp în Doncaster şi a avut o revelaţie spontană, hotărându-se să-şi dedice viaţa explicării logicii gândirii omeneşti.21 în urma acelei plimbări a rezultat algebra booleană, care constituie o manieră de a simplifica expresiile logice folosind simboluri şi numere (de ex., „şi”, „său” şi „nu”). Astfel, de exemplu, dacă calculăm „adevărat *şi* adevărat”, rezultatul ar fi adevărat, ceea ce ar corespunde comutatoarelor unui calculator. Boole a avut nevoie de două decenii pentru a-şi pune pe hârtie teoriile. Şi a mai fost nevoie de încă o sută de ani pentru ca cineva să-şi dea seama că logica şi probabilitatea booleană ar putea ajuta calculatoarele să treacă de la automatizarea unor calcule aritmetice către a deveni maşini de gândire mai complexe. Nu exista o cale de a construi o maşină gânditoare – procesele, materialele şi energia nu erau încă disponibile-aşadar teoria nu a putut fi verificată.

Saltul de la maşinile de gândire teoretică la computere care să înceapă să imite gândirea omenească s-a produs în anii 1930, odată cu apariţia a doua articole revoluţionare: „A Symbolic Analysis of Switching and Relay Circuits” de Claude Shannon respectiv „On Computable Numbers, With an Application to the *Entscheidungsproblem*” de Alan Turing. Ca student în inginerie electrică la MIT, Shannon a făcut un curs opţional de filosofie o abatere neobişnuită. *Legile gândirii* ale lui Boole au devenit principala sursă pentru teza lui Shannon. Îndrumătorul lui, Vannevar Bush, l-a încurajat să aplice logica booleană circuitelor fizice. Bush construise o variantă avansată a Motorului Analitic al lui Lovelace şi Babbage – prototipul lui se numea Analizorul Diferenţial – iar structura era întrucâtva improvizată. La momentul respectiv nu exista o teorie sistematică prin care să se

impună cum ar trebui legate circuitele electrice. Descoperirea lui Shannon a constat în transpunerea logicii simbolice a lui Boole asupra circuitelor electrice, apoi în explicarea felului în care logica booleană ar putea fi folosită pentru crearea unui circuit în care să se adauge 1 şi O. Shannon şi-a dat seama că pentru un calculator e nevoie de două straturi: cel fizic (recipientul) şi cel logic (codul).

În timp ce Shannon încerca să aplice logica booleană unor circuite fizice, Turing testa maşina universală de traducere a lui Leibniz care ar putea reprezenta toată cunoaşterea matematică ştiinţifică. Turing urmărea să dovedească ceea ce la vremea respectivă se numea *Entscheidungsproblem*, sau „problema alegerii”. Pe scurt, problema poate fi rezumată astfel: nu există niciun algoritm care să stabilească dacă o afirmaţie matematică arbitrară este adevărată sau falsă. Răspunsul era negativ. Turing a putut dovedi că nu există niciun algoritm, însă, ca produs secundar, el a descoperit un model matematic al unei maşini de calcul care ar putea fi folosită în orice domeniu.22

Şi de aici s-a schimbat totul. Turing şi-a dat seama că un program şi datele acestuia ar putea fi stocate într-un calculator – din nou, în anii ’30 această idee era una radicală. Până în acel moment, toată lumea era de acord că maşina, programul şi datele erau independente. Maşina universală a lui Turing a reuşit să explice pentru prima dată de ce cele trei sunt interconectate. Din punct de vedere mecanic, logica aflată la baza circuitelor şi comutatoarelor putea de asemenea să fie stocată sub forma programului sau a datelor. Gândiţi-vă la importanţa acestor teorii. Recipientul, programul şi datele constituiau o singură entitate, asemănătoare oamenilor. Şi noi suntem recipiente (corpurile noastre), programe (funcţii celulare autonome) şi date (ADN-ul nostru, combinat cu informaţiile senzoriale directe şi indirecte).

Între timp, lunga tradiţie a automatelor care

începuse cu 400 de ani în urmă cu un mic călugăr care se ruga s-a intersectat în final cu Shannon şi Turing. Compania americană Westinghouse a construit un robot denumit „Elektro Moto-Omul Westinghouse” pentru Expoziţia Universală din 1939. Era un gigant rudimentar, colorat auriu, şi cu roţi la picioare. Avea 48 de circuite care funcţionau pe platforma unei centrale telefonice. Elektro răspundea, pe baza unor mesaje preînregistrate pe un patefon, la comenzile vocale date printr-un receptor de telefon. Era un computer antropomorf, capabil de decizii rudimentare – spre exemplu ce să spună fără o implicare directă, în timp real a unui om.

Judecând după titlurile ziarelor, nuvelele SF şi buletinele de ştiri din acea perioadă, e clar că oamenii au fost luaţi prin surprindere, şocaţi şi preocupaţi de această evoluţie. Pentru ei, părea că începuse peste noapte epoca „maşinilor gânditoare”. Scriitorul SF Isaac Asimov a publicat „Liar!”, o povestire pătrunzătoare în numărul din mai 1941 al revistei *Astounding Science Fiction*. Era o reacţie la studiile avangardiste de care devenise conştient, formulând în cadrul acesteia cele Trei Legi ale Roboticii:

1. Un robot nu poate răni un om, sau prin inacţiune să permită ca un om să sufere.

2. Un robot trebuie să asculte ordinele pe care i le dau oamenii, mai puţin în cazurile în care aceste ordine s-ar afla în conflict cu Prima Lege.

3. Un robot trebuie să-şi protejeze existenţa atât timp cât această protecţie nu intră în conflict cu Prima sau a Doua Lege.

Ulterior, Asimov a adăugat o lege „Zero”, care să le supună pe toate celelalte: „Un robot nu poate face rău omenirii, sau prin inacţiune, să permită ca omenirea să sufere”.

Dar ce ar gândi o maşină gânditoare?

În 1943, cercetătorii în psihiatrie Warren McCulloch şi Walter Pitts de la University of Chicago au publicat un articol important „Calculul logic al ideilor imanente

activităţii nervoase” prin care descriau un nou tip de sistem care modelează neuronii biologici într-o arhitectură neurală de reţea pentru inteligenţă. Dacă recipientele, programele şi datele sunt interconectate, aşa cum susţinea Turing, şi dacă şi oamenii ar fi nişte recipiente la fel de elegant create şi capabile de procesarea datelor, atunci rezultă că e posibilă construirea unei maşini de gândit dacă aceasta ar fi modelată folosind organul responsabil pentru gândirecreierele noastre. Ei au propus o teorie a minţii şi creierului bazată pe calcul, o „reţea neurală”. În loc să ne concentrăm asupra maşinii ca hardware şi a programului ca software, ei şi-au imaginat un nou tip de sistem simbiotic capabil să înghită cantităţi mare de date, aşa cum facem şi noi oamenii. Computerele nu erau încă suficient de puternice la acea vreme pentru a putea testa această teorie, însă articolul lor i-a inspirat pe alţii să încerce să construiască un nou sistem inteligent de calcul.

Legătura dintre sistemele inteligente de calcul şi procesul decizional autonom a devenit mai clară odată ce savantul John von Neumann, specialist în informatică, fizică şi matematică a publicat un tratat impresionant de matematici aplicate. Scrisă împreună cu economistul de la Princeton Oskar Morgenstern în 1944, carte de 641 de pagini explica cu lux de amănunte felul în care teoria jocului stătea la baza tuturor deciziilor economice. Această lucrare a dus la colaborarea dintre von Neumann şi armata americană, care dezvolta la acel moment un nou tip de calculator electric denumit Electronic Numerical Integrator and Computer, sau ENLAC, pe scurt/ La început, instrucţiunile ENIAC erau înserate fizic în sistem, ceea ce însemna că fiecare reprogramare presupunea recrearea fizică a conexiunilor. Inspirat de Turing, McCulloch şi Pitts, von Neumann a elaborat o metodă de a stoca programele în calculator. Aceasta a reprezentat trecerea de la prima eră a calculatoarelor (tabularea) către o nouă epocă a sistemelor programabile.

a Integrator Numeric Electronic şi Calculator (n.t.)

Turing însuşi lucra la o reţea neurală alcătuită din calculatoare cu programe stocate pe sistemele respective. În 1949, publicaţia *The London Times* relata afirmaţia lui Turing: „Nu văd de ce maşina nu ar putea concura în vreunul dintre domeniile care sunt în prezent apanajul intelectului omenesc şi să nu ajungă eventual să concureze de pe poziţii de egalitate. Nu cred că puteţi trage linie când vine vorba de sonete, deşi poate că această comparaţie e uşor nedreaptă pentru că un sonet scris de o maşină va fi mai bine apreciat de o altă maşină”. După un an, într-un articol publicat în jurnalul de filosofie *Mind*, Turing vorbea despre întrebările ridicate de Hobbes, Descartes, Hume şi Leibniz. În acest articol, el a propus o teză şi un test: dacă într-o bună zi, un calculator va deveni capabil să răspundă la întrebări de o manieră care să nu poată fi distinsă de cea omenească, atunci acel calculator va „gândi”. Probabil că aţi auzit de acest articol sub o altă denumire: testul Turing.

Articolul începe cu o întrebare devenită celebră, o întrebare pe care şi-au pus-o şi la care au răspuns atâţia filosofi, teologi, matematicieni şi oameni de ştiinţă înaintea lui: „Poate o maşină să gândească?” Turing însă, conştient de dezbatere veche de secole despre minte şi maşini a respins întrebarea, susţinând că e prea vastă pentru a genera o discuţie fertilă. „Maşină” şi „a gândi” sunt termeni ambigui, care lasă prea mult loc interpretărilor subiective. (La urma urmei, existau deja patru secole în care se scriseseră cărţi şi articole despre înţelesul acestor cuvinte.)

Jocul presupunea o amăgire, iar calculatorul „câştiga” dacă putea trece drept un om. Testul se desfăşoară în felul următor: există un om, o maşină, şi o persoană care pune întrebări într-o cameră separată. Obiectivul jocului este ca interogatorul să-şi dea seama care dintre răspunsuri provin de la o persoană, şi care provin de la maşină. La începutul jocului, interogatorul primeşte etichete, X şi Y – neştiind însă care este cea care

corespunde calculatorului şi i se permite să pună doar întrebări de genul „Poate X să-mi spună dacă X joacă şah?” La finalul acestui joc, interogatorul trebuie să-şi dea seama cine era X şi cine era Y. Cealaltă persoană are misiunea de a-l ajuta pe interogator să identifice maşina, iar maşina trebuie să păcălească interogatorul să creadă că e de fapt om. În privinţa jocului, Turing a scris:

Cred că peste vreo 50 de ani va deveni posibil să programăm nişte calculatoare cu o capacitate de stocare de I09, să le punem să joace jocul imitaţiei atât de bine încât un interogator mediu nu ar avea mai mult de 70% şanse să identifice corect personajele după cinci minute de întrebări.23

Turing era însă un om de ştiinţă şi el ştia că această teorie nu putea fi dovedită, cel puţin nu în cursul vieţii lui. Problema s-a dovedit a fi însă nu faptul că Turing nu dispunea de dovezi empirice care să dovedească că maşinile ar putea într-o zi să gândească, şi nici momentul – Turing a spus că ar fi nevoie să aşteptăm până la finalul secolului XX pentru a putea desfăşura acest test. „Sperăm ca maşinile să poată concura în cele din urmă cu oamenii în toate domeniile intelectuale”, scria Turing.

Adevărata problemă era să facem în aşa fel încât să ajungem credem că maşinile ar putea într-o bună zi să vadă, să raţioneze şi să îşi aducă aminte – şi că oamenii ar putea încerca să împiedice acest progres. Astfel, colegii lui ar fi trebuit să observe cogniţia lipsiţi de spiritualism şi să creadă în plauzibilul unor maşini inteligente care, spre deosebire de oameni, să ia decizii într-un mod inconştient.

Vara şi iarna inteligenţei artificiale

Profesorii Marvin Minsky (matematică şi neurologie) şi John McCarthy (matematică) împreună cu Claude Shannon (matematician şi criptografia Beli Labs) şi Nathaniel Rochester (informatician la IBM) au propus în 1955 un atelier de două luni dedicat operei lui Turing şi posibilităţii ca maşinile să poată învăţa. Teoria lor era următoarea: dacă era posibil să descriem fiecare

caracteristică a inteligenţei umane, atunci o maşină ar putea fi învăţată să o simuleze.24 Era însă nevoie de un grup mare, divers de experţi din multe domenii diferite. Ei considerau că puteau înregistra progrese semnificative strângând laolaltă un grup interdisciplinar de cercetători şi lucrând intensiv, fără întreruperi, pe durata verii.

Selecţia grupului avea o importanţă critică. El avea să devină reţeaua de ingineri, sociologi, informaticieni, psihologi, matematicieni, fizicieni şi specialişti în cogniţie care urmau să pună întrebările fundamentale şi să le răspundă, întrebări despre ce înseamnă să „gândim”, cum funcţionează „minţile” noastre şi cum putem face maşinile să înveţe la fel cum o facem noi oamenii. Se intenţiona ca această reţea diversă să continue să colaboreze în domeniul cercetării şi să ducă acest nou domeniu mai departe. Pentru că era vorba de o abordare nouă, interdisciplinară cu obiectivul de a construi maşini care să gândească, aceştia aveau nevoie de un nou nume prin care să-şi descrie activitatea. Au găsit o descriere ambiguă dar elegantă: inteligenţa artificială.

McCarthy a creat o listă preliminară cu 47 de experţi de care credea că avea nevoie pentru a pune bazele grupului şi ale activităţii de cercetare. A fost un proces tensionat prin care s-au stabilit toate persoanele importante care trebuiau să fie prezente la demersul teoretic şi practic al inteligenţei artificiale. Minsky era în mod special preocupat de faptul că la întâlnire aveau să lipsească două voci esenţiale-Turing, care decedase cu doi ani în urmă, şi von Neumann, suferind de cancer în fază terminală.25

Cu toate eforturile de a aduna un grup divers cu cea mai bună combinaţie de aptitudini complementare, ei aveau şi un minus. Toţi erau albi, deşi existau mulţi oameni de culoare care lucrau în domeniile pe care McCarthy şi Minsky voiau să le reunească Toţi erau angajaţi ai marilor companii din domeniul tehnologiei de la vremea respectivă (IBM, Beli Labs) sau proveneau de la o mână de facultăţi.

Chiar dacă erau numeroase femei de geniu care contribuiseră semnificativ în domeniul ingineriei, informaticii, matematicii sau fizicii, ele au fost excluse.26 Toţi participanţii erau bărbaţi, în afară de Gloria, soţia lui Marvin Minsky. Fără să-şi dea seama de propriile prejudecăţi, aceşti oameni de ştiinţă care sperau să înţeleagă cum funcţionează mintea omenească, cum gândim şi cum ar putea maşinile să înveţe de la toţi oamenii şi-au limitat drastic datele la cei care arată şi vorbesc la fel ca ei.

În anul următor, grupul s-a adunat la ultimul etaj al facultăţii de matematică din Dartmouth pentru a studia teoria complexităţii, simularea limbajului natural, reţelele neurale, legătura dintre aleatoriu şi creativitate şi maşinile care pot învăţa. În timpul săptămânii se întâlneau în sala de curs pentru o discuţie generală înainte de a se dedica unor sarcini concrete. Profesorii Allen Newell, Herbert Simon şi Cliff Shaw au descoperit o metodă de a descoperi dovezi ale teoremelor logice şi simulau procesul – un program pe care l-au denumit „Teoreticianul Logic” – într-una din sesiunile generale. Acesta a fost primul program care a mimat aptitudinile omeneşti de rezolvare a problemelor. (Programul a dovedit într-un final 38 din cele 52 de teoreme din *Principia Mathematica* a lui Alffed North Whitehead şi Bertrand Russell, o carte de referinţă despre bazele matematicii.) Claude Shannon, care sugerase cu câţiva ani în urmă că ar trebui să învăţăm calculatoarele să joace şah împotriva oamenilor, a avut ocazia să prezinte un prototip al acestui program, aflat încă în faza de dezvoltare.27

Aşteptările lui McCarthy şi Minsky de a progresa spectaculos în domeniul inteligenţei artificiale nu s-au materializat în vara aceea la Dartmouth. Nu era suficient timp – ca să nu mai vorbim de puterea de calcul – pentru ca inteligenţa artificială să treacă de la teorie la practică.28 Cu toate acestea, în vara aceea au fost stabilite trei practici care au stat la baza inteligenţei artificiale aşa

cum avea ea să se manifeste ulterior:

1. Inteligenţa artificială avea să fie concepută, construită, testată şi perfecţionată de o colaborare între marile companii din domeniul informaticii şi cercetătorii academici.

2. Progresul inteligenţei artificiale este o activitate costisitoare, prin urmare fiind necesară comercializarea activităţii – fie colaborând prin parteneriate cu agenţiile guvernamentale sau cu armata, sau prin construirea unor produse şi sisteme care să poată fi vândute.

3. Cercetarea şi construirea inteligenţei artificiale se bazează pe o reţea interdisciplinară de cercetători, ceea ce presupune înfiinţarea de la zero a unui nou domeniu academic. Acest lucru a însemnat şi că persoanele din acest domeniu au avut tendinţa de a recruta oameni pe care-i cunoşteau deja, păstrând omogenitatea reţelei dar limitându-i perspectiva asupra lumii.

Tot în acea vară a mai avut loc un eveniment interesant. În timp ce grupul dezbătea întrebarea lui Turing – *Poate o maşină să gândească*? aceştia nu au putut cădea de acord asupra modalităţii optime de a dovedi răspunsul, respectiv de a construi o maşină care să poată învăţa. Unii membri preferau o abordare biologică. Ei credeau că se pot folosi reţele neurale care să poată imprima inteligenţei artificiale un bun simţ şi un raţionament logic – că ar fi posibil ca maşinile să fie în general inteligente. Alţii susţineau că ar fi imposibil ca să se creeze vreodată o copie atât de fidelă a structurilor omeneşti de gândire. Ei preferau în schimb o abordare inginerească. În loc să scrie comenzi pentru a rezolva probleme, un program ar putea ajuta sistemul să „înveţe” dintr-un set de date. Sistemul ar face predicţii pe baza datelor, iar un supervizor uman ar verifica acele răspunsuri – perfecţionând maşina pe parcurs. Privită din acest unghi, capacitatea maşinilor de a „învăţa” primea o definiţie îngustă, care se rezuma la învăţarea unei sarcini concrete, cum ar fi a juca dame.

Psihologul Frank Rosenblatt, prezent la atelierul de la Dartmouth, voia să modeleze felul în care creierul uman procesează informaţiile vizuale şi, prin urmare, cum învaţă să recunoască obiecte. Pe baza studiilor din acea vară, Rosenblatt a creat un sistem denumit Perception. El intenţiona să construiască un simplu program cadru care să reacţioneze la feedback. Aceasta a fost prima reţea neurală artificială care opera creând conexiuni între diferite elemente de procesare într-o structură stratificată. Fiecare neuron mecanic primea multe semnale diferite, apoi folosea un sistem de ponderare matematică pentru a stabili ce semnal să genereze. În această structură paralelă puteau fi accesaţi mai multe procesoare simultanlucru care însemna nu doar că era rapidă, ci că putea procesa continuu mai multe date.

Iată care este importanţa acestei structuri: deşi ea nu însemna neapărat că un calculator ar putea „gândi”, ea ne demonstra însă cum putem face un calculator să înveţe. Noi oamenii învăţăm încercând şi dând greş. Pentru a cânta gama de major la pian trebuie să apăsăm pe clape în ordinea corectă. La început, degetele, urechile şi ochii noştri nu au memorat tiparul corect însă dacă exersăm – dacă repetăm gama constant, corectându-ne de fiecare dată în cele din urmă vom reuşi. Când am luat lecţii de pian şi am cântat prost gamele, profesorul mă corecta, dar dacă le nimeream, primeam un abţibild. Abţibildul mă ajuta să înţeleg că luasem deciziile corecte atunci când cântam. Şi reţeaua neurală a lui Rosenblatt funcţionează la fel. Sistemul învăţa cum să optimizeze răspunsul repetând de mii de ori aceleaşi funcţii şi reţinea ce a învăţat aplicând această cunoaştere pentru noi probleme. El antrena sistemul folosind o tehnică denumită retropropagare. În faza iniţială, un om evaluează dacă reţeaua neurală artificială lua decizia corectă. Dacă lua decizia corectă, procesul era consolidat. Dacă nu, se ajustau parametrii şi se aplică un alt test.

În anii de după atelier s-au înregistrat progrese

semnificative în domeniul unor probleme complexe pentru oameni, cum ar fi folosirea inteligenţei artificiale pentru rezolvarea teoremelor matematice. Cu toate acestea, o sarcină simplă – cum ar fi recunoaşterea limbajului – rămânea o provocare uriaşă, fără soluţii imediate. Înainte de cercetarea în domeniul inteligenţei artificiale, mintea fusese considerată dintotdeauna o cutie neagră. Datele intrau, şi din cutie ieşea un răspuns fără să existe o modalitate de a observa acest proces. Primii filosofi, matematicieni şi oameni de ştiinţă spuneau că acesta era rezultatul unei creaţii divine. Oamenii de ştiinţă din epoca modernă ştiau că era rezultatul a sute de mii de ani de evoluţie. Abia în anii ’50, după vara de la Dartmouth, cercetătorii au început să creadă că pot decoda (cel puţin pe hârtie) cutia neagră şi să observe procesele cognitive, apoi să înveţe calculatoarele să imite comportamentul nostru de răspuns la stimuli.

Până în acest punct, calculatoarele fuseseră doar un instrument de automatizare a calculelor. Prima eră a informaticii, marcată de maşini care puteau să calculeze numere, lăsa loc celei de-a doua ere, cea a computerelor programabile. Acestea erau sisteme mai rapide şi mai compacte care dispuneau de suficientă memorie pentru a stoca seturi de instrucţiuni. Programele puteau acum şi stocate local, şi, mai ales, puteau fi scrise în limba engleză, şi nu sub forma unui cod complicat pentru maşini. Devenea clar că nu aveam nevoie de automate sau de recipiente antropomorfe pentru ca aplicaţiile inteligenţei artificiale să devină utile. Inteligenţa artificială putea să se găsească într-o cutie simplă, fără trăsături omeneşti, şi cu toate acestea să fie extrem de utilă.

Atelierul de la Dartmouth l-a inspirat pe matematicianul britanic I.J. Good să scrie despre o „maşină a hiperinteligenţei” care ar putea crea maşini mai bune decât noi. Prin urmare ar rezulta o „explozie a inteligenţei, iar inteligenţa omenească ar fi lăsată mult în urmă. Aşadar, această primă maşină ultra inteligentă ar fi ultima invenţie

pe care ar trebui să o mai facă omul”. 29

În toată această evoluţie şi-a făcut intrarea şi o femeie, cel puţin cu numele. Informaticianul de la MIT Joseph Weizenbaum a conceput un sistem denumit ELIZA, un program de conversaţie denumit după personajul din piesa *Pygmalion* de George Bernard Shaw.30 Această evoluţie a fost importantă pentru reţelele neurale şi inteligenţa artificială pentru că a fost una din primele încercări de a procesa limbajul natural, iar programul accesa diferite scenarii create în prealabil pentru a conversa cu oameni în carne şi oase. Cel mai faimos dintre acestea se numea DOCTOR31 şi imita un psiholog empatic folosind recunoaşterea tiparelor pentru a oferi răspunsuri frapant de umane.

Atelierul de la Dartmouth atrăsese acum atenţia comunităţii internaţionale, ca şi cercetătorii care s-au trezit pe neaşteptate în lumina reflectoarelor. Erau nişte staruri rock educate, care le ofereau oamenilor obişnuiţi o perspectivă fantastică, nouă asupra viitorului. Vă aduceţi aminte de Rosenblatt, psihologul care crease prima reţea neurală? El a declarat pentru *Chicago Tribune* că în curând maşinile nu vor avea doar programe de tip ELIZA, capabile să ofere câteva sute de răspunsuri – ci vor putea să participe la şedinţe şi să transcrie „ca orice secretară”. El a promis nu doar că va crea cel mai mare „aparat de gândit” din lume, ci şi că îl va produce în timp de câteva luni.32

Şi ce susţineau Simon şi Newell, autorii Teoreticianului Logic? Ei au început să emită previziuni îndrăzneţe, radicale despre inteligenţa artificială, spunând că în zece ani – *adică până în 1967* – calculatoarele vor:

bate toţi campionii de şah şi vor deveni campioane mondiale descoperi şi dovedi o nouă teoremă matematică importantă scrie o muzică pe care şi cei mai aprigi critici să o aprecieze33

Între timp, Minsky făcea previziuni despre o maşină inteligentă care să poată face mult mai multe decât să

scrie după dictare, să joace şah sau să scrie muzică. El considera că până la finalul vieţii sale, maşinile vor dobândi inteligenţa artificială generală – cu alte cuvinte, că vor deveni capabile de gândire complexă, se vor putea exprima prin limbaj şi vor putea face alegeri.34

Cercetătorii de la atelierul din Dartmouth au scris articole şi cărţi. Au dat interviuri la radio, televiziune, în ziare şi reviste. Ştiinţa era însă dificil de explicat aşa că în mod frecvent explicaţiile au fost distorsionate şi cuvintele lor au fost scoase din context. Lăsând la o parte previziunile exagerate, aşteptările publicului în privinţa inteligenţei artificiale au devenit din ce în ce mai fantastice, şi asta şi pentru că povestea a fost relatată greşit. Spre exemplu, Minsky ar fi spus conform revistei *Life*: „În trei-opt ani vom avea o maşină de nivelul general de inteligenţă al unui om mediu. Vreau să spun că vom avea o maşină care va putea să citească din Shakespeare, să schimbe uleiul la maşină, să bârfească la muncă, să spună o glumă sau să se bată”. 35 în acelaşi articol, jurnalistul se referă la Alan Turing denumindu-l „Ronald Turing”. Minsky, deşi entuziast, nu voia să spună că mai era doar un pas până la roboţi vorbitori ci mai degrabă oferise un răspuns arogant. În absenţa unui context şi a explicaţiilor, percepţia publică asupra inteligenţei artificiale a început să se distorsioneze.

Faptul că în 1968 Arthur Clarke şi Stanley Kubrick au decis să facă un film despre un viitor în care maşinile vor avea nivelul general de inteligenţă al omului mediu nu a ajutat. Ei voiau să spună o poveste despre oameni şi maşinile de gândit – iar Minsky a fost cooptat în calitate de consultant. Dacă încă nu aţi ghicit, e vorba de un film pe care îl cunoaşteţi deja – 2001: Odiseea spaţială, axat în jurul unei inteligenţe artificiale pe nume HAL 9.000 care a învăţat simţul umorului şi creativitatea de la cei care l-au creat şi ameninţă să ucidă pe oricine ar vrea să-l deconecteze. Unul dintre personaje, Victor Kamiński, a fost denumit în cinstea lui Minsky.

Putem spune că la mijlocul anilor ’60, inteligenţa artificială devenise parte a spiritului vremurilor, şi toată lumea fetişiza viitorul. Crescuseră şi aşteptările legate de succesul comercial al inteligenţei artificiale datorită unui articol apărut într-o publicaţie obscură a industriei radio. Cu un titlu simplist ca „Cum să îndesăm mai multe componente în circuitele integrate”, articolul, scris de co-fondatorul intel, Gordon Moore, postula o teorie conform căreia numărul de tranzistori care pot încăpea pe o placă de circuite la acelaşi preţ se va dubla la fiecare 18 – 24 de luni. Această idee îndrăzneaţă a devenit cunoscută drept Legea lui Moore, iar la început, această teză părea să fie corectă. Computerele deveneau din ce în ce mai puternice, şi capabile de o multitudine de exerciţii, nu doar probleme de matematică. Comunitatea inteligenţei artificiale urma să primească un sprijin, pentru că asta presupunea că teoriile lor urmau curând să poată fi verificate. Legea lui Moore mai introducea şi posibilitatea extraordinară ca procesoarele fabricate de oameni să depăşească într-o bună zi puterile minţii omeneşti, care are o capacitate de înmagazinare limitată biologic.

Toate aceste exagerări şi acest articol au canalizat multe fonduri în direcţia inteligenţei artificiale – chiar dacă persoanele care nu făceau parte din reţeaua de la Dartmouth nu înţelegeau exact ce este inteligenţa artificială. Nu existau încă produse şi nici metode practice de a aplica reţelele neurale la scară mare sau tehnologia necesară. Fiindcă oamenii credeau acum în *posibilitatea* maşinilor gânditoare, a fost suficient pentru a obţine fonduri substanţiale din mediul privat dar şi de la guvern.

Spre exemplu, guvernul Statelor Unite a finanţat un program ambiţios în care inteligenţa artificială trebuia folosită pentru a traduce limbajul. Era în toiul Războiului Rece, iar guvernul voia să aibă la dispoziţie un sistem de traducere instant a limbii ruse care să ofere o eficienţă mai mare, reduceri de costuri şi sporirea preciziei. Se părea că programarea ar fi putut oferi soluţia. Dintr-o colaborare

între Institutul de Literatură şi Lingvistică de la Universitatea Georgetown şi IBM a rezultat un prototip de aparat de traducere din rusă în engleză care avea un vocabular limitat la 250 de cuvinte şi era specializat doar în chimie organică. Demonstraţia cu public a provocat un entuziasm iar traducerea automatizată a ajuns pe prima pagină a ziarului *New York* Timesdar şi a altor 6 ziare.

Fondurile erau abundente – graţie finanţării agenţiilor guvernamentale, universităţilor şi marilor companii din domeniul tehnologiei – şi se părea pentru o vreme că nimeni nu ţinea socoteala. Cu toate acestea, dincolo de articole şi prototipuri, inteligenţa artificială nu reuşea să ofere ceea ce se prevăzuse şi promisese. S-a dovedit că progresele semnificative reprezentau o provocare mult mai mare decât anticipaseră pionierii modernităţii.

La scurt timp după aceea, s-a cerut o analiză a felului în care inteligenţa artificială ar putea fi folosită în viaţa de zi şi a implementării ei concrete. National Academy of Sciences a înfiinţat o comisie consultativă la solicitarea National Science Foundation, a Departamentului Apărării şi CIA. Comisia a constatat că părerile erau contradictorii în ce priveşte posibilitatea unui program bazat pe inteligenţă artificială care să efectueze traduceri în timp real, concluzionând că „nu există traduceri efectuate de aparate ale unor texte ştiinţifice generale, şi nici perspective imediate pentru ca ele să apară în viitor”. 36 Un raport produs ulterior pentru British Science Research Council susţinea că cercetătorii au exagerat avansul tehnologic în domeniul inteligenţei artificiale, oferind o prognoză pesimistă pentru toate ariile de bază de cercetare din acest domeniu. James Lighthill, un matematician de la Cambridge, era principalul autor al raportului; reproşul cel mai puternic era că aceste tehnici de debut ale inteligenţei artificiale – spre exemplu, a învăţa un calculator să joace dame – nu puteau să producă soluţii pentru problemele mai complexe ale lumii reale.37

În urma acestor rapoarte, politicienii din Statele Unite şi Marea Britanie au cerut răspunsuri la o nouă întrebare: de ce finanţăm extravaganţele unor teoreticieni? Guvernul Statelor Unite, inclusiv DARPA, a retras finanţarea pentru proiectele de automatizare a traducerii. Companiile şi-au schimbat prioritătile de la programe cronofage de cercetare generală a inteligenţei artificiale la programe de amplitudine mai mică menite să rezolve probleme concrete. Dacă în primii ani de după atelierul de la Dartmouth a existat o perioadă de mari aşteptări şi optimism, deceniile care au urmat acelor rapoarte au devenit cunoscute drept iarna inteligenţei artificiale. Finanţarea a fost tăiată, studenţii au preferat alte domenii de cercetare şi avansurile au fost oprite.

Chiar şi McCarthy şi-a revizuit negativ predicţiile. „Oamenii pot face aceste lucruri foarte uşor pentru că aşa sunt construiţi”, spunea McCarthy.38 E mult mai greu să înţelegem care sunt modalităţile prin care recunoaştem vorbire – procesele fizice şi cognitive care stau la baza recunoaşterii limbajului. McCarthy dădea exemplul unei colivii pentru a ilustra provocările din domeniul inteligenţei artificiale. Să spunem că v-aş cere să-mi construiţi o colivie, şi nu v-aş oferi niciun fel de alte date. Probabil că aţi construi o cutie dotată cu un vârf, un fiind, şi nişte laturi. Dacă v-aş mai da o informaţie – că pasărea e un pinguin – atunci poate nu i-aţi mai pune un vârf. Prin urmare, prezenţa vârfului depinde de câţiva parametri: de informaţiile pe care vi le dau, şi de toate asocierile pe care vi le provoacă cuvântul „pasăre”, cum ar fi că majoritatea păsările sunt zburătoare. Avem cu toţii un anumit context şi nişte presupuneri implicite. Ca inteligenţa artificială să poată reacţiona la fel ar presupune ca ea să dispună de mult mai multe informaţii şi instrucţiuni explicite.39 Iarna inteligenţei artificiale urma să dureze trei decenii.40

Ce a urmat: învăţarea jocurilor

Deşi finanţarea s-a redus, mulţi dintre cercetătorii de la Dartmouth au continuat să lucreze cu inteligenţa

artificială – şi au continuat să predea materia unor noi studenţi. Între timp, legea lui Moore a rămas valabilă, iar calculatoarele au devenit din ce în ce mai puternice.

În anii ’80, unii cercetători şi-au dat seama cum ar putea vinde anumite trăsături ale inteligenţei artificiale – iar acum exista suficientă putere de calcul şi o reţea de cercetători în continuă creştere care constatau că din munca lor se pot face bani. Astfel, interesul publicului a fost repornit, şi odată cu el, şi investiţiile în inteligenţa artificială au fost reluate. Japonia a anunţat în 1981 un plan pe 10 ani de dezvoltare a inteligenţei artificiale intitulat „Generaţia a cincea”. Guvernul Statelor Unite a reacţionat înfiinţând Microelectronics and Computer Technology Corporation, un consorţiu de cercetare menit să asigure competitivitatea economiei naţionale, în Marea Britanie, fondurile care fuseseră tăiate în urma raportului lui James Lighthill au fost realocate. Între 1980 şi 1988, industria inteligenţei artificiale a crescut de la câteva milioane de dolari la câteva miliarde.

Calculatoarele mai rapide, încărcate cu memorie, puteau acum procesa mai eficient datele, iar accentul era pus pe reproducerea procesului decizional al experţilor, în loc de a construi maşini universale, ca fictivul computer HAL 9.000. Aceste sisteme se axau în principal pe folosirea reţelelor neurale pentru sarcini mărunte, cum ar fi jocurile. În anii ’90 şi 2000 au avut loc nişte succese rezonante. În 1994, un program denumit CHINOOK a jucat şase jocuri de dame împotriva campionului mondial Marlon Tinsley (toate partidele încheindu-se cu remize). CHINOOK a câştigat prin abandonul lui Tinsley.41 în 1997, supercalculâtorul IBM Deep Blue a câştigat în faţa campionului mondial Garry Kasparov, care a cedat stresului unei finale de şase meciuri împotriva unui adversar aparent invincibil. În 2004, Ken Jennings a câştigat în ciuda probabilităţilor 74 de partide consecutive la *Rişti şi câştigi*, stabilind un record mondial la acea vreme pentru sumele încasate dintr-un concurs televizat.

Aşa că atunci când a acceptat provocarea celor de la IBM de a juca împotriva lui Watson, în 2011, se simţea încrezător că va câştiga. Făcuse câteva cursuri de inteligenţă artificială şi presupunea că tehnologia nu era suficient de avansată pentru a-şi da seama de context, subtilităţi semantice şi jocuri de cuvinte. Watson l-a umilit pe Jennings, care a început să-şi piardă încrederea încă de la începutul jocului.

Ceea ce ştiam în 2011 e că inteligenţa artificială e mai bună decât oamenii la anumite sarcini de gândire pentru că poate accesa şi procesa cantităţi uriaşe de informaţii fără să cadă pradă stresului. Inteligenţa artificială poate defini stresul, dar nu are un sistem endocrin pe cap.

Cu toate acestea, jocul străvechi Go reprezenta apogeul cercetătorilor în inteligenţa artificială pentru că putea fi jucat doar pe baza unei strategii convenţionale. Go este un joc străvechi apărut acum peste 3.000 de ani, jucat cu reguli suficient de simple: doi jucători plasează pe rând pietre albe sau negre pe o grilă goală. Pietrele pot fi capturate atunci când sunt înconjurate de culoarea opusă, sau când nu mai există alte spaţii libere sau „libertăţi”. Obiectivul este să acoperiţi teritoriu pe grila de joc, dar asta presupune aptitudini psihologice şi o înţelegere precisă a stării de spirit a adversarului.

Dimensiunea tradiţională a grilei de Go este de 19 cu 19 pătrăţele. Spre deosebire de alte jocuri, precum şahul, pietrele de Go au valori egale. Există 181 de piese negre şi 180 de piese albe (negrul mută întotdeauna primul, de unde diferenţa). În şahunde piesele au diferite valori – jucătorul alb are la dispoziţie 20 de mutări, şi cel negru 20 de mutări. După prima mutare, sunt 400 de poziţii diferite ale tablei. În Go există însă 361 de prime mutări posibile, câte una la intersecţia unei grile complet goale. După primele mutări ale fiecărui jucător, sunt 128.960 de mutări posibile, în total, există I0170 configuraţii posibile – ca să vă daţi seama, e un număr mai mare decât numărul

atomilor din universul cunoscut. Date fiind atâtea poziţii şi mişcări potenţiale, nu există un manual, la fel ca în cazul şahului sau al damelor. Marii maeştri ai jocului de Go se bazează pe scenarii: dacă adversarul mută într-un anume punct, atunci care sunt rezultatele posibile, plauzibile şi probabile, date fiind personalitatea, răbdarea şi starea generală a adversarului?

La fel ca şahul, Go este un joc determinist cu informaţii perfecte, în care nu există niciun element ascuns sau de hazard. Pentru a câştiga, jucătorii trebuie să-şi echilibreze emoţiile şi trebuie să stăpânească la perfecţie arta subtilităţii. În şah se pot anticipa eventualele mişcări probabile ale unui jucător; o tură se poate mişca doar vertical sau orizontal pe tablă. Numărul mişcărilor potenţiale al acelei piese este deci limitat. Prin urmare, e mai uşor să înţelegeţi cine câştigă un joc de şah cu mult timp înainte ca vreo piesă să fi fost capturată sau regele să fie pus în şah. În Go lucrurile nu stau aşa. Uneori e nevoie de un mare maestru în Go care să-şi dea seama ce se petrece şi cine pare la un moment dat să câştige. Complexitatea jocului de Go i-a asigurat un public printre împăraţi, matematicieni şi fizicieni. Acesta este motivul pentru care cercetătorii inteligenţei artificiale au fost dintotdeauna fascinaţi de jocul de Go şi au încercat să dezvolte programe care să joace Go.

Jocul de Go s-a dovedit dintotdeauna o nucă tare pentru cercetătorii inteligenţei artificiale. Deşi un calculator poate fi programat să înveţe regulile jocului, cum rămâne cu regulile care ar permite înţelegerea caracteristicilor omeneşti ale adversarului? Nimeni nu mai construie un algoritm suficient de puternic care să poată face faţă complexităţilor acestui joc. În 1971, un program scris de informaticianul Jon Ryder a încercat o abordare tehnică, dar a pierdut în faţa unui începător. În 1987, un program mai puternic denumit Nemesis a concurat pentru prima dată într-un turneu împotriva unui om. În 1994, programul Go Intellect se dovedea un jucător competent.

Chiar şi în ciuda unui handicap semnificativ, acesta a pierdut toate cele trei partide – împotriva unor copii. În toate aceste cazuri, calculatoarele făceau fie mişcări de neînţeles, fie jucau prea agresiv, sau calculau greşit strategia adversarului.

În cadrul acestor demersuri erau şi câţiva cercetători care reluau ideea reţelelor neurale, o idee lansată de Marvin Minsky şi Frank Rosenblatt la prima întâlnire de la Dartmouth. Cercetătorul cognitivist Geoff Hinton şi informaticienii Yann Lecun şi Yoshua Bengio credeau că reţelele neurale vor avea aplicaţii practice concrete – cum ar fi detectarea automată a fraudelor cu cărţi de credit sau recunoaşterea automată a caracterelor optice, pentru a citi documente sau cecuri – dar şi că ele vor pune bazele viitorului inteligenţei artificiale.

Hinton – profesor la universitatea din Toronto-a fost cel care şi-a imaginat un nou tip de reţea neurală, alcătuită din diverse straturi care extrag diferite informaţii până când reţeaua recunoştea ce anume căuta. Singura modalitate de a oferi unui sistem de inteligenţă artificială acest gen de cunoaştere, credea el, era de a dezvolta algoritmi de învăţare care să le permită calculatoarelor să înveţe pe cont propriu. În loc să le învăţăm să facă o singură treabă simplă deosebit de bine, reţelele urmau să fie construite pentru a se învăţa singure.

Aceste reţele neurale noi, profunde urmau să necesite un nou tip de învăţare pentru maşini – învăţarea profundă – pentru a antrena calculatoarele să efectueze sarcini omeneşti însă cu mai puţină (sau chiar nicio) supraveghere din partea oamenilor. Unul dintre avantajele imediate era faptul că proiectul putea fi extins la o scară mare. Într-o reţea neurală, câţiva neuroni fac câteva alegeri-numărul alegerilor posibile creşte exponenţial cu fiecare strat. Reformulând: oamenii învaţă individual, dar omenirea învaţă colectiv. Imaginaţi-vă o mare reţea neurală profundă, care învaţă ca un întreg – având posibilitatea de a-şi spori cu timpul viteza, eficacitatea şi

randamentul.

Un alt avantaj îl reprezintă faptul că aceste sisteme ar fi libere să înveţe pe cont propriu, fără să fie limitate de capacităţile cognitive şi imaginaţia oamenilor. Creierul omenesc are nişte praguri chimice şi metabolice care limitează puterea de procesare a calculatoarelor umede din capul nostru. Nu putem evolua pe cont propriu în mod semnificativ, iar ritmul evoluţiei nu se potriveşte cu aspiraţiile noastre tehnologice actuale. Ni se oferea promisiunea că învăţarea profundă ar accelera şi însăşi evoluţia inteligenţei, care ar presupune doar temporar un efort din partea oamenilor.

O reţea neurală profundă urma să primească un set elementar de parametri despre date din partea unei persoane-iar sistemul ar urma apoi să înveţe pe cont propriu recunoscând tipare şi folosind mai multe straturi de procesare. Pentru cercetători, atracţia învăţării profunde este că maşinile sunt programate să ia decizii în mod imprevizibil. Gândirea într-o manieră pe care noi oamenii nu am putut-o concepe – sau pune în practică – e esenţială atunci când încercăm să rezolvăm marile probleme, care nu au avut niciodată soluţii clare.

Comunitatea inteligenţei artificiale a considerat că reţelele neurale profunde sunt nişte elucubraţii ale unui cercetător excentric, îndoiala lor s-a acutizat atunci când şi-au dat seama că din cauză că procesele de învăţare profundă se petrec în paralel, cercetătorii inteligenţei artificiale nu le-ar putea observa în timp real. Cineva ar trebui să construiască sistemul, apoi să aibă încredere că lua deciziile corecte.

A câştiga şi a pierde

Hinton a continuat să lucreze, organizând ateliere pe tema acestui concept cu studenţii săi, dar şi cu Lecun şi Bengio, şi a publicat articole, începând cu 2006. În anul 2009, laboratorul lui Hinton aplica reţelele neurale profunde pentru recunoaşterea vocală, şi dintr-o întâlnire întâmplătoare cu Li Deng, un cercetător de la Microsoft, a

rezultat o etapă pilot a acestei tehnologii. Deng, un specialist chinez în învăţare profundă, era unul dintre pionierii recunoaşteriii vocale folosind învăţarea profundă la scară largă. Tehnica intrase în teste la Google în 2010. Doar doi ani mai târziu, reţelele neurale profunde erau folosite în produsele comerciale. Dacă folosiţi Google Voice şi serviciile de transcriere vocală, ele se bazează pe învăţare profundă, iar tehnica a devenit baza tuturor asistenţilor digitali pe care-i folosim astăzi. Siri, Google şi Alexa de la Amazon sunt toate exemple bazate pe învăţarea profundă. Comunitatea interdisciplinară de cercetători în inteligenţă artificială a crescut semnificativ de la vara din Dartmouth. Aceste trei practici esenţiale au rămas însă valabile-colaborarea dintre companiile mari din domeniul tehnologiei şi cercetătorii din mediul academic; succesul comercial ca promotor al progreselor în domeniul inteligenţei artificiale şi omogenitatea reţelei de cercetători.

Toate aceste progrese americane nu au trecut neobservate la Beijing. China are acum un propriu ecosistem incipient, dar în creştere, iar guvernul central sprijină cercetătorii care vor să-şi publice lucrările. Numărul de articole dedicate inteligenţei artificiale publicate de cercetători chinezi a crescut cu peste 100% între 2010 şi 2017.*42* La drept vorbind, articolele şi brevetele nu înseamnă automat că aceste studii vor conduce şi la o întrebuinţare comună a acestei tehnologii, însă statisticile ne arată cât de tulburaţi au fost liderii chinezi de avansul occidentalilormai ales când vine vorba de jocul de Go.

În ianuarie 2014, Google a început să investească sume considerabile în inteligenţa artificială, printre care se numără o sumă de peste 500 de milioane de dolari pentru achiziţia unei firme noi înfiinţate în domeniul învăţării profunde denumită Deep Mind, fondată de neurobiologul Demis Hassabis, un fost mare jucător de şah în copilărie, cercetătorul în domeniul învăţării maşinilor Shane Legg şi

antreprenorul Mustafa Suleyman. Faptul că dezvoltaseră un program denumit Alpha Go a constituit o parte din atracţia acestei echipe.

Într-o perioadă de câteva luni au fost gata să testeze programul Alpha Go împotriva unui jucător uman. S-a programat un meci între Deep Mind şi Fan Hui, un jucător chinez profesionist de Go, şi unul dintre cei mai puternici maeştri din Europa. Deoarece a juca Go pe calculator nu e tocmai la fel cu a juca pe o tablă fizică, s-a decis ca unul dintre inginerii Deep Mind să mute pe tablă în locul calculatorului şi să-i transmită calculatorului mişcările acestuia.

Înaintea jocului, Toby Manning, unul dintre capii British Go Association, a jucat o rundă de test împotriva lui Alpha Go – şi a pierdut la o diferenţă de 17 puncte. Manning a făcut nişte greşeli, dar şi programul a greşit. Acestuia i-a trecut un gând bizar prin cap: dacă Alpha Go jucase conservator? Era cu putinţă ca programul să fi jucat doar suficient de agresiv încât să-l înfrângă pe Manning, în loc să-l strivească?

Jucătorii s-au aşezat la o masă, Fan Hui purtând o cămaşă în dungi şi o jachetă de piele maro, Manning în centru şi inginerul de cealaltă parte. Jocul a început. Hui a deschis o sticlă de apă şi s-a uitat la tablă. Jucând cu piesele de culoare neagră, trebuia să înceapă. În primlee 50 de mişcări jocul a fost liniştit – Hui încerca în mod evident să sondeze punctele forte şi cele slabe ale lui Alpha Go. Unul dintre primele semne: programul nu juca agresiv decât dacă rămânea în urmă. A fost un meci strâns. Alpha Go a câştigat o victorie foarte strânsă, de doar 1, 5 puncte.

Hui s-a folosit de această informaţie când a început a doua partidă. Dacă Alpha Go nu avea de gând să joace agresiv, atunci Hui s-a decis că va lupta de la început. Alpha Go a început să mute mai repede. Hui a amintit că poate avea nevoie de mai mult timp ca să se gândească la mutări. La mutarea 147, Hui a încercat să împiedice ca

Alpha Go să obţină un teritoriu mare în mijlocul tablei dar mutarea a dat greş, şi a fost forţat să renunţe.

În jocul al treilea, Hui a mutat mai agresiv, şi Alpha Go a procedat la fel. Undeva pe la jumătatea meciului, Hui a făcut o greşeală catastrofală pe care Alpha Go a speculat-o, apoi o altă greşeală mare care a condus la finalul jocului. Plin de frustrare, Hui a trebuit să se scuze şi a plecat să se plimbe pentru a-şi putea redobândi calmul şi a termina partida. Din nou, stresul a doborât o minte strălucită în vreme ce inteligenţa artificială şi-a văzut fără scrupule de obiectiv.

Alpha Go – un program bazat pe inteligenţă artificială a bătut cu 5-0 un jucător profesionist de Go. A reuşit să câştige analizând mult mai puţine mutări decât computerul Deepblue al IBM. Când Alpha Go a bătut un jucător uman, acesta nu ştia că joacă un joc, sau ce înseamnă un joc, sau de ce oamenilor le face plăcere să joace jocuri.

Hanjin Lee, un jucător bine cotat de Go din Korea, a analizat jocul dintre cei doi. „Impresia mea generală e că Alpha Go părea un jucător mai bun decât Fan, dar nu mi-am putut da seama cât de mare e acea diferenţă… poate că devine mai puternic atunci când joacă cu un adversar mai puternic”. 43

Accentul pus pe jocuri-a bate un om într-o confruntare directă – defineşte succesul folosind un set relativ redus de parametri. Ajungem astfel la întrebare filosofică pentru epoca modernă a inteligenţei artificiale. Pentru ca sistemele inteligenţei artificiale să câştige – pentru a îndeplini obiectivele pe care le-am definite nevoie ca oamenii să piardă, atât dintr-o perspectivă trivială, cât şi una mai profundă?

Alpha Go a continuat să joace, depăşind orice adversar ca un maer maestru şi demoralizând comunitatea jucătorilor profesionişti de Go. După ce a câştigat împotriva campionului mondial cu 3-0, Deep Mind a anunţat retragerea din competiţie a sistemului, spunând că echipa se va concentra asupra unor noi provocări.44

Echipa a încercat apoi să caute o variantă prin care Alpha Go să treacă de la un sistem puternic care putea fi antrenat să-i bată pe cei mai buni jucători la un sistem care ar putea învăţa singur să ajungă la acelaşi nivel de expertiză fără să recurgă la ajutorul oamenilor.

Prima variantă a Alpha Go se baza pe intervenţia oamenilor şi un set iniţial de date format din 100.000 de jocuri de Go pentru ca sistemul să poată învăţa să joace. Următoarea generaţie a sistemului a fost construită să înveţe de la zero. La fel ca un jucător începător, această variantă – denumită Alpha Go Zerova trebui să înveţe pe cont propriu de la zero, fără să dispună de variante din care să poată alege sau să aibă măcar o definiţie a acţiunilor pe care le presupun mutările. Sistemul nu ar lua doar decizii – rezultatul unor calcule care pot fi programate în mod explicit – ci ar face alegeri, ceea ce necesită o judecată.45 Aceasta presupunea că designerii Deep Mind aveau o putere extraordinară, chiar dacă nu îşi dau seama de asta. Zero avea să înveţe condiţiile, valorile şi motivările pentru deciziile şi alegerile pe care le face în timpul jocului de la programatori.

Zero a concurat împotriva sa, ajustând şi perfecţionând de unul singur procesele decizionale. Fiecare joc începea cu câteva mişcări aleatorii, şi cu fiecare câştig, Zero îşi actualiza sistemul şi juca apoi luând în calcul ceea ce învăţase. Zero a avut nevoie de doar 70 de ore de joc pentru a ajunge la aceeaşi putere pe care o avea Alpha Go când câştigase împotriva celor mai buni jucători din lume.46

După care s-a produs ceva interesant. Echipa Deep Mind a aplicat tehnica unei instalări separate a lui Alpha Go Zero folosind o reţea mai mare şi i-a permis să se antreneze şi să joace singură time de 40 de zile. Programul nu doar că a reuşit să capete toate cunoştinţele despre Go adunate de oameni de-a lungul istoriei, ci a reuşit să învingă varianta avansată a Alpha Go în 90% din jocuriasta înseamnă că Zero devenise un student mai bun decât cei

mai mari maeştri de go şi un profesor mai bun decât cei care l-au antrenat, şi nu înţelegem în întregime ce a făcut pentru a deveni atât de inteligent.47 Cam cât de inteligent credeţi că e? Ei bine, puterea unui jucător de Go se măsoară pe baza unui index denumit Elo, care stabileşte probabilitatea de a pierde sau a câştiga pe baza performanţelor anterioare. Marii maeştri şi campionii mondiali au tendinţa să aibă un scor în jur de 3500. Zero are un scor de peste 5.000. Prin comparaţie, aceste genii, campionii mondiali joacă ca nişte amatori, şi este statistic improbabil ca orice jucător uman să poată vreodată învinge acest sistem bazat pe inteligenţa artificială.

Ştim care este condiţia care a permis acest tip de învăţare. Nefolosind vreo expertiză umană sau date, creatorii Zero au îndepărtat constrângerile omeneşti puse inteligenţei artificiale. Se pare că influenţa oamenilor ar fi făcut sistemul mai puţin performant. Reuşita a constat în crearea unui sistem care avea capacitatea de a gândi într-o manieră nouă şi de a face propriile alegeri.48 Sistemul a reprezentat un salt brusc, neaşteptat, care prefigurează un viitor în care sistemele de inteligenţă artificială ar putea analiza radiografiile pacienţilor de cancer, evalua datele climatologice şi analiza sărăcia într-o metodă neomenească – putând eventual să conducă la progrese pe care cercetătorii umani să nu le fi putut niciodată obţine pe cont propriu.

În timp ce Zero juca împotriva sa, a descoperit strategii de Go pe care oamenii le-au dezvoltat de-a lungul a mii de ani – ceea ce înseamnă că a învăţat să gândească la fel ca oamenii care l-au conceput. La început făcea aceleaşi greşeli, detecta aceleaşi tipare şi variaţii şi tipare şi întâmpina aceleaşi obstacole ca noi. Odată ce Zero a devenit suficient de puternic, a abandonat mişcările pe care le-ar fi făcut oamenii şi şi-a dezvoltat propriile preferinţe.49 Odată ce Zero a început să devină autonom, el şi-a dezvoltat strategii creative pe care nimeni nu le mai văzuse, sugerând că poate maşinile *gândesc deja* în moduri

care sunt atât recognoscibile, cât şi străine de noi.

Zero a mai dovedit şi că algoritmii sunt capabili să înveţe ajutorul oamenilor, şi că intervenţia oamenilor e cea care înfrâna dezvoltarea inteligenţei artificiale. Aceasta înseamnă că am putea încredinţa maşinilor problemele pe care noi înşine nu le putem prevedea sau rezolva.

În decembrie 2017, echipa Deep Mind a publicat un articol în care arăta că Zero a devenit capabil să înveţe – nu doar Go, dar şi alte informaţii. Zero a început să joace alte jocuri, cum ar fi şah şi shogi (un joc japonez asemănător cu şahul), care sunt mai puţin complexe dar necesită totuşi creativitate şi strategie. Zero învăţa însă mult mai repede ca înainte. El a reuşit să dezvolte o putere supraomenească, dene în teles, în mai puţin de 24 de ore de joc. Hassabis a început apoi să aplice tehnicile pe care le-a dezvoltat în echipa sa pentru proiectul Zero pentru a construi o „maşină de învăţat cu scop general”, un set de algoritmi care se pot adapta şi care imită sistemele noastre biologice şi pot fi antrenate. În loc să umple sistemele inteligenţei artificiale cu cantităţi uriaşe de informaţii şi instrucţiuni referitoare la felul în care pot fi accesate, echipa vă învăţa maşinile cum să înveţe. Spre deosebire de oameni care pot să obosească, să se plictisească sau să-şi piardă atenţia atunci când studiază, maşinile vor urmări cu orice preţ un obiectiv.

Acesta a reprezentat un moment de cotitură în lunga istorie a inteligenţei artificiale din mai multe motive. Mai întâi, sistemul se comporta de o manieră imprevizibilă, luând decizii care nu păreau raţionale pentru creatorii lui. Sistemul a reuşit să înfrângă un adversar uman fără ca să se poată înţelege sau reproduce complet cum a putut face asta. Apariţia lui Zero prefigurează un viitor în care inteligenţa artificială şi-ar putea construi propriile reţele neurale şi să obţină o cunoaştere pe care nu am putea s-o înţelegem, în al doilea rând, apariţia lui a cimentat cele două parcursuri paralele pe care le urmează evoluţia inteligenţei artificiale în prezent: alarmată, China alocă

bani şi oameni pentru a-şi face propriile produse mai competitive, în timp ce în Statele Unite ne aşteptăm să vedem produse fantastice bazate pe inteligenţa artificială şi care să fie disponibile în curând. Viabilitatea reţelelor neurale profunde şi a învăţării profunde explică spatele frenezia din zilele noastre legate de inteligenţa artificială-ca să nu mai amintim de creşterea spectaculoasă a finanţării în Statele Unite şi a programelor naţionale ale Chinei pentru viitor.

Ca o companie din cadrul Alphabet (compania mamă a Google), Deep Mind are 700 de angajaţi, iar o parte dintre ei au misiunea de a dezvolta cât mai repede produse care să poată fi comercializate. În martie, divizia de servicii cloud a Google a anunţat că vinde un serviciu de citire cu voce tare a unui text bazat pe tehnologia Deep Mind cu 16 dolari per 1 milion de caractere de text procesate.50 Unul dintre anunţurile de senzaţie de la conferinţa Google I/O din 2018 a fost Duplex, un asistent vocal care va efectua automat apeluri în numele clienţilor şi va discuta cu personalul pentru a face rezervări la restaurant sau programări la coafor, cu tot cu „mm”-uri şi „ah”-uri. La baza acelui produs stă Wave Net, un program de generare software prin inteligenţă artificială, parte a portofoliului Deep Mind.51

Între timp, cercetătorii din domeniul inteligenţei artificiale din altă divizia a Alphabet denumită Google Brain au dezvăluit că deţineau o inteligenţă artificială capabilă să genereze propriile inteligenţe artificiale. (Aţi înţeles?) Sistemul, denumit Autorul, automatizează crearea modelelor de învăţare automată printr-o tehnică denumită consolidarea învăţării. Autorul funcţionează ca un soi de „părinte” – o reţea neurală profundă care decide să creeze reţele de inteligenţă artificială „copii” pentru anumite sarcini concrete, ţintite. Fără să i se spună să facă asta, Autorul a generat un copil denumit NASNet şi l-a învăţat să recunoască obiecte cum ar fi maşinile, semafoarele, oamenii sau genţile în videoclipuri. Neîmpovărat de ego,

îndoieli sau lipsa încredenitrăsături pe care chiar şi cei mai sclipitori informaticieni le au – NASNet a reuşit o acurateţe de 82, 7% în recunoaştere imaginilor. Cu alte cuvinte, sistemul a reuşit o performanţă mai bună decât cea a programatorilor umani-inclusiv a celor care au creat sistemul părinte.52

Echipele de programatori care creează aceste sisteme menite să ia decizii şi să facă alegeri sunt conduse de bărbaţi. Grupul e uşor mai diversificat decât cercetătorii de la Dartmouth din cauza unei evoluţii semnificative ulterioare-intrarea Chinei în joc. În ultimii ani, China a devenit un centru important de cercetare în inteligenţa artificială, datorită unor demersuri colosale de finanţare ale guvernului la universităţi dar şi companii ca Baidu, Alibaba şi Tencent.

Baidu a reuşi o performanţă chiar mai mare decât a lui Zero: să transfere aptitudini dintr-un domeniu în altul. Pentru oameni e o misiune simplă, însă nu şi pentru inteligenţa artificială. Baidu a urmărit să depăşească acest obstacol învăţând o reţea neurală profundă să navigheze într-o lume virtuală 2D folosind doar limbajul natural, aşa cum vorbesc părinţii cu copiii. Agentul inteligenţei artificiale Baidu primea comenzi cum ar fi „te rog du-te către măr” sau „poţi să te deplasezi către grila dintre măr şi banană?” apoi primea recompense pentru acţiunile corecte. Pare uşor, însă gândiţi-vă care sunt implicaţiile: la finalul experimentului nu doar că inteligenţa artificială a Baidu devenise capabilă să înţeleagă un limbaj despre care nu ştiuse nimic, dar sistemul a şi învăţat ce este o grilă bi-dimensională, că se poate deplasa de-a lungul ei, cum să se deplaseze, că există mere şi banane, şi cum să le poată diferenţia.

La începutul acestui capitol, am lansat patru întrebări: *Pot maşinile să gândească*? *Ce ar însemna ca o maşină să „gândească”*? *Ce înseamnă pentru voi, dragi cititori, să gândiţi*? *Cum aţi putea să vă daţi seama că gândiţi gânduri originale*? Acum după ce aţi aflat istoria

îndelungată a acestor întrebări, micul grup de oameni care au pus bazele inteligenţei artificiale şi principalele practici care mai sunt folosite, aş dori să vă ofer câteva răspunsuri.

Da, maşinile pot gândi. Capacitatea de a trece un test conversaţional, cum ar fi testul Turing sau recent apăruta Schemă Winograd, propusă de Hector Levesque în 2011 şi care se concentrează asupra raţionamentelor elementare, provocând inteligenţa artificială să răspundă la o întrebare care conţine pronume ambigue, nu măsoară neapărat capacitatea unui sistem de inteligenţă artificială în alte domenii, ci dovedeşte doar că acest sistem poate gândi într-un cadru lingvistic, asemenea oamenilor.53 Suntem cu toţii de acord că Einstein era un geniu, chiar dacă metodele acceptate prin care se măsura inteligenţa în vremea respectivă – cum ar fi examenele – indicau un alt rezultat. Einstein gândea într-o manieră pe care profesorii lui nu o puteau înţelegeei au presupus desigur că nu era inteligent. De fapt însă, la vremea respectivă nu exista o modalitate relevantă de a măsura inteligenţa lui Einstein. Acelaşi lucru este valabil şi pentru inteligenţa artificială.

Maşinile care gândesc pot lua decizii şi alege într-un mod care să aibă impact în lumea reală, iar pentru a face asta au nevoie de un scop şi un obiectiv. În cele din urmă, şi-ar putea dezvolta o judecată. Acestea sunt atribute ale sufletului, conform filosofilor şi teologilor. Fiecare suflet este o manifestare a viziunii şi voinţei divine, fiind creat şi înzestrat de un creator. Şi maşinile care gândesc au creatori, ei sunt noii zei ai inteligenţei artificiale, în majoritate bărbaţi care locuiesc cu preponderenţă în Statele Unite, Europa occidentală şi China şi care au legături cu cei nouă titani. Sufletul inteligenţei artificiale reprezintă o manifestare a viziunii şi voinţei lor pentru viitor.

Şi da, pentru a concluziona, maşinile care gândesc sunt capabile de gândire originală. După ce învaţă prin experienţă, ar putea constata că există o soluţie diferită, sau că se impune o nouă clasificare. Inteligenţa artificială

nu trebuie să inventeze o nouă formă de artă pentru a ne demonstra creativitatea.

Aceasta înseamnă că există în aceste maşini o minte. Această minte e încă tânără şi în curs de formare, şi e probabil ca ea să evolueze pe un parcurs pe care noi să nu-l înţelegem. În capitolul următor vom vorbi despre elementele acestei minţi, valorile celor Nouă Titani şi consecinţele sociale, politice şi economice neprevăzute ale acestei revoluţii a inteligenţei artificiale.

CAPITOLUL DOI

LUMEA INSULARĂ A TRIBURILOR INTELIGENTEI ARTIFICIALE

Demersurile seculare de a construi o maşină capabilă să gândească nu au dat roade decât recent. Deşi aceste maşini ar putea lăsa impresia că „gândesc”, trebuie să înţelegem de la capul locului că ele nu gândesc cu siguranţă la fel ca *noi*.

Viitorul inteligenţei artificiale este construit de o minoritate relativ redusă de oameni din cadrul unor grupuri mici, atomizate. Repet, eu cred că aceşti oameni au intenţii bune. Date fiind însă aceste grupuri insulare care colaborează strâns unele cu celelalte, prejudecăţile lor inconştiente şi o oarecare lipsă de perspectivă vor avea tendinţa să devină cu timpul noi sisteme de credinţă şi comportamente acceptate. Ceea ce în trecut ar fi putut părea neobişnuit, chiar greşit, devine un nou normal cotidian, iar *această* gândire este programată în aceste maşini.

Cei care lucrează în domeniul inteligenţei artificiale fac parte dintr-un soi de trib. Ei sunt nişte oameni care trăiesc şi lucrează în America de nord şi China. Sunt oameni care au absolvit aceleaşi universităţi, şi respectă aceleaşi reguli sociale. Triburile sunt foarte omogene. Vorbim de oameni bine educaţi şi înstăriţi. În mare parte triburile se compun din bărbaţi, conduşi aproape în unanimitate de către bărbaţi-preşedinţii, membrii

consiliilor de administraţie şi directorii sunt aproape toţi bărbaţi. Omogenitatea reprezintă o problemă şi în China, unde membrii tribului sunt cu preponderenţă chinezi.

Problema triburilor reprezintă şi unul dintre motivele care le conferă putere. În grupurile insulare, prejudecăţile cognitive se amplifică şi consolidează, instalându-se în subconştient. Prejudecăţile cognitive sunt un substitut al gândirii raţionale, care este mai înceată şi presupune un consum mai mare de energie. Cu cât un trib devine mai interconectat şi mai respectat, cu atât gândirea de grup şi comportamentul grupului par mai normale. Veţi constata mai târziu că merită să reţineţi această concluzie.

Care este activitatea triburilor din domeniul inteligenţei artificiale? Aceste triburi construiesc sisteme de inteligenţa artificială îngustă, capabilă să efectueze o singură sarcină la un nivel comparabil sau superior cu al oamenilor. Implementările comerciale ale aplicaţiilor de inteligenţă artificială îngustă – prin extensie, triburile înseleiau deja decizii în numele nostru în căsuţele noastre de e-mail, când căutăm pe internet, când facem poze cu telefonul, când conducem maşina sau când solicităm un card de credit sau un împrumut. Triburile lucrează deja la următoarea etapă: sisteme de inteligenţă artificială generală care vor îndeplini sarcini cognitive mai vaste pentru că sunt maşini create să gândească, la fel ca noi. Cine sunt însă aceşti „noi” care servesc drept model acestor sisteme? Ce fel de valori, idealuri şi perspective asupra lumii sunt transmise sistemelor?

Pe scurt, nu ale voastre, şi nici ale mele. Inteligenţa artificială e concepută după chipul şi asemănarea tribului, acordând prioritate valorilor, idealurilor şi perspectivelor asupra lumii ale celor care au creat-o. Ea începe însă să gândească şi pe cont propriu.

Liderii tribali

Tribul inteligenţei artificiale are un strigăt de luptă cunoscut, antrenant: *dă greş repede, şi dă greş des*. O altă variantă – „mişcă-te repede şi distruge lucruri” era motto-

ul oficial al Facebook până nu demult. Ideea de a comite greşeli şi a accepta eşecul se află într-un contrast radical cu filosofia corporaţiilor uriaşe ale Americii, care evită riscurile şi se mişcă cu viteza melcului, constituind un obiectiv lăudabil. Tehnologiile complicate precum inteligenţa artificială presupun experimente şi ocazia de a da greş în mod repetat, cu scopul de a nimeri varianta corectă. Există, însă, şi un revers al medaliei. Această mantră face parte dintr-o ideologie larg răspândită între cei Nouă Titani: *mai întâi construieşte*, *cere-ţi iertare mai târziu*.

În ultima vreme am tot auzit scuze. Facebook şi-a cerut scuze pentru rezultatul relaţiei cu Cambridge Analytica. În timpul desfăşurării scandalului, Facebook a anunţat în luna septembrie 2018 ca un atac a vulnerabilizat datele personale a peste 50 de milioane de utilizatori, una dintre cele mai însemnate breşe de securitate din istoria digitală. Se pare însă că directorii au decis să nu anunţe imediat utilizatorii.54 La doar o lună după aceea, Facebook a anunţat lansarea serviciului Portal, un ecran de video-conferinţe care să rivalizeze cu Echo Show al Amazon, renunţând la promisiunile legate de securitatea datelor personale pe care le făcuse anterior. La început, Facebook spusese că nu va folosi serviciul Portal pentru a colecta date personale în vederea ţintirii anunţurilor publicitare către utilizatori. După ce jurnaliştii au cerut informaţii suplimentare, compania s-a văzut nevoită să ofere o lămurire bizară: deşi Portal nu va folosi datele voastre pentru a afişa reclame, datele strânse în timp ce folosiţi serviciul-persoanele apelate, melodiile ascultate pe Spotifyar putea fi folosite pentru a vă oferi reclame Facebook prin alte servicii şi reţele.55

În luna aprilie 2016, Jeff Dean, şeful de proiect al Google Brain scria că în timpul unei sesiuni de întrebări pe Reddit, compania exclusese femeile şi persoanele de culoare. Nu era un gest premeditat ci o omisiune involuntară, şi cred cu tărie că nu a fost voit, însă

organizatorii pur şi simplu nu s-au gând cum ar putea diversifica publicul sesiunii.

Dean spunea că pune preţ pe diversitate şi că Google va trebui să-şi îmbunătăţească practicile.56

„Unul dintre lucrurile care chiar mi-a plăcut la programul nostru Brain Residency era că rezidenţii provin dintr-o gamă largă de profiluri, domenii de expertiză (avem de exemplu fizicieni, matematicieni, biologi, neurobiologi, ingineri electricieni dar şi informaticieni) şi diferă între ei şi în alte moduri, contribuind la demersurile noastre de cercetare”, scria Dean. „Din experienţa mea, ori de câte ori aduci laolaltă oameni cu expertize, perspective diferite etc vei ajunge să reuşești mult mai mult decât ai fi putut pe cont propriu pentru că nu există un singur om înzestrat complet cu aptitudinile şi perspectiva necesară pentru un asemenea exerciţiu”. 57

Google a publicat în luna iunie 2018 un raport al diversităţii care cuprindea pentru prima oară date despre numărul de angajaţi în funcţie de categorie. În cadrul raportului, Google spunea că la nivel global forţa lor de muncă se compune în proporţie de 69, 1% din bărbaţi. În Statele Unite, doar 2, 5% dintre angajaţi sunt negri, în vreme ce 3, 6% sunt hispanici sau latino-americani. Cu toate afirmaţiile îndrăzneţe ale Google privind nevoia de a avea diversitate, aceste procentaje – deja scăzutenu demonstrau vreo schimbare faţă de anii precedenţi, în 2014 proporţia negrilor fiind de 2% în vreme ce hispanicii şi latino-americanii reprezentau 3% din total.58

Spre meritul lor, Google au lansat în ultimii ani o iniţiativă împotriva prejudecăţilor inconştiente cu ateliere şi cursuri menite să ajute angajaţii companiei să înveţe mai multe despre stereotipurile sociale şi atitudinile generale care s-ar fi putut forma inconştient privind genul, rasa, aspectul, vârsta, educaţia, politica sau bunăstarea. Unii angajaţi ai Google sunt de părere că aceste cursuri au mai degrabă un caracter de umplutură, iar o angajată de culoare a explicat că în cadrul cursului s-a vorbit cu

precădere despre „relaţii interpersonale şi sentimente rănite în loc să se discute despre discriminare şi inegalitate, ceea ce transmite mesajul către angajaţi că diversitatea e doar, o rubrică de bifat”. 59

Cu toate acestea, în timp ce avea loc acest curs, Google recompensa comportamentele nesănătoase ale conducerii. Andy Rubin, creatorul emblematicului sistem mobil de operare Android, fusese invitat să demisioneze după ce o angajată susţinuse că acesta ar fi obligat-o să întreţină relaţii sexuale orale. Google i-a plătit lui Rubin suma de 90 de milioane de dolari pentru a demisiona în tranşe lunare de 2, 5 milioane de dolari în primii 2 ani şi 1, 5 milioane lunar în următorii doi ani. Directorul diviziei de cercetare şi dezvoltare X, Richard Devaul, a hărţuit sexual o femeie în cursul unui interviu de angajare, spunându-i că el şi soţia lui au o căsătorie fără angajamente, insistând apoi să mângâie acea candidată pe spate la un festival tehno. În mod evident, ea nu a fost angajată. I s-a cerut să prezinte scuze dar nu să demisioneze. Un vice preşedinte care se ocupa de divizia de căutare a Google a avut probleme atunci când o angajată l-a acuzat că ar fi pipăit-oacuzaţia a fost considerată credibilă, aşa că a fost concediat însă cu un pachet de compensare de mai multe milioane de dolari, între 2016 şi 2018, Google a renunţat la 13 directori din motive de hărţuire sexuală.60

Această reacţie surprinde impactul şters pe care îl au multe programe de conştientizare în rândul firmelor din domeniul tehnologiei şi al companiilor care le finanţează. Motivul este următorul: deşi oamenii ar putea să fie mai conştienţi de propriile prejudecăţi după curs, ei nu devin neapărat motivaţi sau stimulaţi să-şi modifice comportamentul.

Atunci când vorbim de lipsa diversităţii în rândul informaticienilor, conversaţia pendulează de regulă între gen şi rasă. Există însă şi alte dimensiuni cărora nu li se acordă la fel de multă atenţie, ca de pildă ideologia politică şi religia. Într-o analiză a Stanford Graduate School of

Business efectuată în 2017, bazată pe răspunsurile a peste 600 de lideri şi antreprenori, se arată că membrii tribului se consideră în proporţie covârşitoare progresişti care susţin Partidul Democrat. În perioada alegerilor din 2016, membrii tribului au sprijinit-o în proporţie covârşitoare pe Hillary Clinton.

Membrii tribului sunt de acord cu impozitarea mai mare a persoanelor înstărite, sprijină avorturile, se opun pedepsei cu moartea, vor un regim mai strict pentru permisele de port armă şi susţin legalizarea căsătoriilor homosexuale.61

Afirmaţia că persoanele aflate la conducerea Google, Apple, Amazon, Facebook, Microsoft sau IBM nu reprezintă în mod fidel o imagine a *tuturor* americanilor s-ar putea aplica şi altor companii din alte domenii. Diferenţa este însă că aceste companii sunt cele care dezvoltă sisteme autonome de luare a deciziilor care sunt menite să reprezinte toate interesele noastre. Criticile provin nu doar din partea femeilor sau a persoanelor de culoare, ci şi din partea unor grupuri neaşteptate: conservatorii şi susţinătorii Partidului Republican. În luna mai 2018, Comitetul Naţional Republican i-a trimis o scrisoare lui Mark Zuckerberg acuzând Facebook de discriminare împotriva americanilor de orientare conservatoare reclamând, printre altele, că: „În ultimii ani au apărut îngrijorări legate de suprimarea mesajelor conservatoare pe Facebook… inclusiv cenzurarea articolelor cu tentă conservatoare… Ne alarmează numeroasele aserţiuni conform cărora Facebook blochează conţinutul publicat de grupuri şi jurnalişti conservatori”. 62 Scrisoarea, semnată de Ronna McDaniel, preşedintele Comitetului Naţional Republicat şi Brad Parscale, şeful de campanie pentru realegerea preşedintelui Trump în 2020, solicită transparenţă în ce priveşte felul în care algoritmii Facebook stabilesc utilizatorii care sunt expuşi la publicitate politică şi analizarea prejudecăţilor împotriva conţinutului şi liderilor conservatori.

Problema e că McDaniel şi Parscale nu greşesc. În timpul alegerilor febrile din 2016, angajaţii Facebook au manipulat cu bună ştiinţă platforma pentru a exclude ştirile cu conţinut conservator – chiar şi prin articole împotriva lui Clinton şi care începuseră să capete propria notorietate. Câţiva dintre „curatorii de ştiri” de la Facebook, aşa cum sunt aceştia cunoscuţi, au afirmat că li s-a cerut să „introducă” anumite articole în flux, chiar dacă acestea nu începuseră să fie urmărite. Curatorii au împiedicat şi afişarea articolelor favorabile unor candidaţi ai Partidului Republican, precum Rând Paul. Echipa de curatori a Facebook era alcătuită dintr-un număr mic de jurnalişti care absolviseră facultăţi private de elită de pe coasta de est, şi, ca să fim sinceri, acest aspect susţine explicaţiile oferite de conservatori vreme de decenii.

În luna august 2018, peste 100 de angajaţi Facebook au folosit un grup intern pentru a dezavua „cultura politică monocromă, intolerantă faţă de alte păreri”. Unul dintre inginerii seniori ai Facebook, Brian Amerige, scria: „Pretindem că suntem deschişi tuturor perspectivelor, dar atacăm – adesea în haită – pe oricine prezintă o perspectivă care pare să se opună ideologiei de stânga”. 63 A vorbi despre diversitate – a cere iertare, a promite că vei face mai bine – nu e totuna cu a rezolva problema diversităţii în bazele de date, algoritmii şi structurile care compun sistemul inteligenţei artificiale. Când vorbele nu conduc la fapte, rezultă un ecosistem de sisteme şi produse care reflectă o anumită prejudecată împotriva oamenilor. În 2016, un robot de securitate controlat de inteligenţa artificială a intrat intenţionat într-un copil în vârstă de 16 luni într-un centru comercial din Silicon Valley.64 Sistemele de inteligenţă artificială aflate la baza jocului video *Elite*: *Dangerous* au creat o serie de super-arme pe care creatorii nu le-au imaginat niciodată, făcând ravagii în joc şi distrugând orice progrese făcute de jucătorii în carne şi oase.65 Există nenumărate probleme legate de siguranţa inteligenţei artificiale, dintre care

unele sunt mari şi evidente: maşinile care conduc singure au trecut deja pe roşu şi chiar au ucis pietoni în câteva rânduri. Aplicaţiile folosite de poliţie etichetează greşit chipurile suspecţilor în mod constant, băgând oameni nevinovaţi în arest. Există nenumărate probleme necunoscute pe care nu le băgăm în seamă pentru că încă nu ne-au afectat personal.

O echipă cu adevărat diversă ar avea o singură trăsătură comună: talentul. Nu ar exista o concentrare a unui anumit gen, sau a unei anumite rase sau etnicităţi. Ar exista reprezentanţi ai mai multor perspective politice şi religioase. Omogenitatea triburilor inteligenţei artificiale constituie o problemă a celor Nouă Titani, însă ea nu începe aici. Problema începe în universităţi, acolo unde se formează triburile.

Triburile apar în cadrul unor medii sociale concentrate în care toată lumea are un scop sau un obiectiv comun, foloseşte acelaşi limbaj şi munceşte într-un ritm relativ similar. În cadrul tribului, grupurile de oameni îşi dezvoltă un simţ al valorilor şi scopurilor împărtăşite. Triburile se formează în locuri precum unităţile militare, gărzile din facultatea de medicină, bucătăriile restaurantelor cu stele Michelin şi organizaţiile studenţeşti. Membrii lor trec împreună prin bune şi rele, reuşite şi eşecuri, fericire şi tristeţe.

Pentru a prelua un exemplu dintr-un domeniu îndepărtat de inteligenţa artificială, în anii ’70 şi ’80, Sam Kinison, Andrew Dice Clay, Jim Carrey, Marc Maron, Robin Williams şi Richard Pryor au locuit împreună într-o casă situată pe Cresthill Road la numărul 8420, la doi paşi de ceea ce avea să devină legendarul club Comedy Store din Los Angeles. Erau doar nişte tineri care locuiau într-o casă, încercând să obţină reprezentaţii pe scenă într-o perioadă în care Bob Hope apărea la televizor cu replici ca „Nu mă gândesc de două ori la femei. În primul gând spun tot”. 66 Acest trib respingea total acest gen de umor, pe care generaţia anterioară îl perfecţionase cu

meticulozitate. Valorile lor erau complet diferite: distrugerea tabuurilor, combaterea nedreptăţilor sociale şi relatarea unor poveşti cu lux de amănunte care aveau tendinţa să prezinte o imagine destul de neplăcută a celor aflaţi în public. Ei repetau între ei. Îşi ţineau de urât după un eşec pe scenă. Experimentau şi învăţau unii de la alţii. Acest trib de comici revoluţionari, inteligenţi a pus bazele viitorului în divertismentul american.67 Acest grup de bărbaţi exercită la nivel colectiv o influenţă chiar şi în zilele noastre.

Inteligenţa artificială a trecut cumva printr-o perioadă asemănătoare de transformări radicale din cauza unui trib din zilele noastre care împărtășește aceleaşi valori, idei şi obiective. Aceştia erau cei trei pionieri ai învăţării profunde amintiţi anterior – Geoff Hinton, Yann Lecun şi Yoshua Bengio – un soi de Sam Kinison şi Richard Pryors ai lumii inteligenţei artificiale în epoca de pionierat a reţelelor neurale profunde. Lecun a studiat cu Hinto la Universitatea din Toronto unde s-au pus bazele unui mic grup de cercetători printre care se numără şi Yoshua Bengio. Ei petreceau perioade lungi de timp împreună, testând idei, teorii şi construind următoarea generaţie a inteligenţei artificiale. „Era o comunitate foarte mică de oameni care anticipaseră că reţelele neurale vor deveni într-o bună zi importante”, spunea Lecun. „Aveam nevoie de o zonă sigură în care să putem organiza ateliere şi întâlniri pentru a ne dezvolta conceptele înainte de a le publica”. 68

Legăturile puternice ale tribului se formează atunci când oameni care lucrează îndeaproape sărbătoresc reuşitele şi suferă în momente grele împreună. Ei îşi dezvoltă un set de experienţe comune, prin care ajung la un limbaj comun, care are ca rezultat un set comun de idei, comportamente şi obiective. Acesta este motivul pentru care atât de multe firme nou înfiinţate, mişcări politice sau simboluri culturale au un început comun: câţiva prieteni care împart un dormitor, o casă sau un

garaj şi lucrează intens la proiecte asemănătoare.

Dacă epicentrele inteligenţei artificiale moderne sunt Silicon Valley, Beijingul, Hangzhou şi Shenzhen, facultăţile sunt esenţa triburilor inteligenţei artificiale. Există doar câteva centre importante. Printre ele, în Statele Unite se numără Carnegie Mellon, Georgia Institute of Technology, Stanford, UC Berkeley, University of Washington, Harvard, Corneli, Duke, MIT, Boston University, McGill University şi Université de Montréal. Aceste universităţi găzduiesc grupuri de cercetare academică, cu puternice legături cu industria IT.

Triburile respectă în general regulile şi ritualurile, aşa că haideţi să vedem în ce constă iniţierea. Procesul începe cu o educaţie temeinică.

În universităţile din America de nord, accentul se pune pe competenţele brute-cum ar fi stăpânirea limbajelor de programare R şi Python, competenţe în procesarea limbajului natural, statistică aplicată şi cunoaşterea teoriei jocurilor, bioinformatică şi vedere digitală. Cursurile din afara tribului, cum ar fi filosofia minţii, literatura feministă islamică sau istoria colonialismului nu sunt văzute cu ochi buni. Dacă încercăm să construim o maşină care să gândească şi să fie capabilă să gândească la fel ca oamenii, pare contraintuitiv să excludem anumite aspecte ale condiţiei umane, în prezent acest gen de cursuri este exclus intenţionat din programa de bază şi e dificil ca ele să fie acceptate chiar şi ca cursuri opţionale.

Tribul pretinde aptitudini, şi în patru ani de facultate sunt multe de învăţat. La Stanford, spre exemplu, studenţii trebuie să facă 50 de ore de matematică, ştiinţe şi inginerie, pe lângă 15 ore de informatică. Deşi în cadrul diplomei se oferă şi un curs de etică, acesta este unul dintre cele cinci cursuri opţionale care pot fi alese pentru a obţine creditele necesare.69 Carnegie Mellon a lansat un program nou de diplomă în inteligenţă artificială în 2018, dând facultăţii un nou imbold şi ocazia să creeze un

program modern de la zero. Ritualurile şi regulile tribului s-au impus însă, iar aptitudinile brute sunt cele care contează. Deşi obţinerea diplomei presupune un curs de etică şi câteva cursuri socio-umane şi de artă, acestea se concentrează în mare parte pe neuorobiologie (psihologie cognitivă, memorie şi cogniţie vizuală), ceea ce e de înţeles dată fiind legătura dintre inteligenţa artificială şi mintea omenească. Nu există cursuri obligatorii care să înveţe studenţii să detecteze prejudecăţile în seturile de date, cum să aplice filosofia în procesul de luare a deciziei sau să le predea etica egalităţii. Aceste cursuri nu recunosc în mod formal că diversitatea socială şi socioeconomică sunt la fel de importante pentru o comunitate ca şi biodiversitatea.

Aptitudinile sunt predate experimentalstudenţii din domeniul inteligenţei artificiale nu-şi petrec timpul citind din cărţi. Pentru a putea învăţa, ei au nevoie de baze de date lexicale, fototeci şi reţele neurale. La un moment dat, una dintre cele mai populare reţele neurale de la universităţi se numea Worzec, fiind construită de echipa Google Brain. Această reţea era un sistem cu două straturi care procesa textul, transformând cuvintele în numere pe care inteligenţa artificială să le poată înţelege.70 Spre exemplu, sistemul a învăţat că „bărbatul e faţă de rege aşa cum femeie e faţă de regină”. Baza de date a hotărât însă şi că „tată e faţă de doctor aşa cum mamă e faţă de asistentă” şi „bărbat e faţă de programator aşa cum femeie e faţă de casnică”. 71 Sistemul la care erau expuşi studenţi avea propriile prejudecăţi. Dacă cineva ar fi vrut să analizeze implicaţiile profunde ale codului cu prejudecăţi, nu existau cursurile care să evidenţieze aceste aspecte.

În 2017 şi 2018 unele universităţi au oferit câteva cursuri noi de etică, ca reacţie la provocările inteligenţei artificiale. Departamentul Berkman Klein Center de la Harvard şi MIT Media Lab oferă împreună un curs nou de etică şi reglementare a inteligenţei artificiale.72 Cursul şi sesiunile sunt excelente, 73 dar nu fac parte din programa

de informatică de bază iar din această cauză materia nu are ocazia să se răspândească şi în alte părţi ale programei.

E drept că etica e o cerinţă a tuturor universităţilor care predau inteligenţa artificială-cerinţa se regăseşte în standardele de acreditare. Pentru a fi acreditate de Accreditation Board for Engineering and Technology, programele de studii în informatică trebuie să arate că studenţii „înţeleg problemele şi răspunderile de ordin profesional, etic, juridic, social şi de siguranţă” şi „au capacitatea de a analiza impactul local şi global al calculatoarelor asupra oamenilor, organizaţiilor şi societăţii”. Vă pot spune însă din experienţă că evaluarea şi măsurarea acestei cerinţe este cel mult subiectivă şi deosebit de greu de pus în practică, mai ales în absenţa unor cursuri obligatorii pentru toţi studenţii. Eu sunt membră a Accrediting Council on Education in Journalism and Mass Communications. Programa de studii a diplomelor de jurnalism şi comunicare pune accentul pe discipline socio-umane, cum ar fi scrisul, reportajele şi producţia media. Cu toate acestea, facultăţile întâmpină în mod regulat probleme în ce priveşte îndeplinirea standardelor noastre legate de probleme şi răspunderi sociale, inclusiv în ce priveşte diversitatea. Facultăţile pot să fie acreditate fără să respecte standardele în ce priveşte diversitatea – şi acest aspect nu e specific doar comisiei de acreditare din care fac parte. Fără o impunere mai strictă a standardelor şi fără demersuri concrete în cadrul universităţilor, cum ar putea o programă încărcată de studiu ca cea a inteligenţei artificiale să rezolve problema?

Facultatea e destul de solicitantă, iar stimulentele oferite noilor angajaţi de către cei Nouă Titani sunt substanţiale. Deşi cursurile opţionale privind literatura africană sau etica funcţionarilor publici ar lărgi cu siguranţă perspectivele asupra lumii ale lucrătorilor din inteligenţa artificială, există o presiune puternică de dezvoltare a ecosistemului. Tribul preferă aptitudinile

legate de muncă astfel încât atunci când proaspeţii absolvenţi intră pe piaţa muncii să poată fi gata de lucru şi să fie membri productivi ai echipelor din care fac parte. De fapt, chiar cursurile care i-ar putea ajuta pe cercetătorii din domeniul inteligenţei artificiale să se gândească mai atent la întreaga omenire le-ar putea diminua şansele de angajare. Cei Nouă Titani folosesc programe bazate pe inteligenţa artificială pentru a filtra CV-urile, căutând anumite cuvinte cheie care descriu aptitudinile. Un portofoliu mai diversificat de cursuri din afara materiei standard ar reprezenta o anomalie sau ar face candidatul mai greu de recrutat.

Faptul că inteligenţa artificială evaluează CV-urile dovedeşte că prejudecăţile nu sunt doar o chestiune de rasă sau gen. Există prejudecăţi şi împotriva filosofiei, literaturii, fizicii sau economiei comportamentale deoarece candidaţii cu multe cursuri opţionale în afara domeniului de bază al inteligenţei artificiale au tendinţa să fie cotaţi mai slab. Sistemul de angajare al tribului, creat să automatizeze munca dificilă de a evalua mii de CV-uri ar putea să excludă aceşti candidaţi care au un profil mai divers şi mai de dorit de studii.

Liderii din domeniul academic vor spune că sunt deschişi la ideea unui curs obligatoriu de etică, chiar şi dacă tribul nu ar dori o programă mai vastă de studiu. (Şi tribul nu vrea acest lucru.) Adăugarea unor cursuri la fel de solicitante din domeniul sociouman, cum ar fi literatura comparată sau religiile lumii, ar forţa scoaterea din program a anumitor cursuri prin care studenţii ar dobândi aptitudinile legate de inteligenţa artificială. Studenţii s-ar revolta că sunt forţaţi să facă şi cursuri aparent inutile, în vreme ce companiile pretind absolvenţilor aptitudini de vârf. Dată fiind concurenţa acerbă de a obţine cei mai buni şi mai inteligenţi studenţi, de ce şi-ar dori vreuna dintre aceste universităţi prestigioase, cum ar fi Carnegie Mellon sau Stanford, să modifice o reţetă de succes?

Domeniul tehnologiei se schimbă mult mai rapid

decât mediul universitar. Un singur curs obligatoriu de etică – gândit în mod special în funcţie de nevoile studenţilor din domeniul inteligenţei artificialenu va fi suficient dacă materia de studiu nu e de actualitate dar mai ales dacă ea nu se regăseşte şi în alte părţi ale programei. Dacă programa nu se poate schimba, ce se poate face în privinţa profesorilor? Ar putea ei să primească răspunderea de a rezolva problema? La scară mare, acest rezultat este improbabil. Profesorii nu sunt încurajaţi să-şi modifice cursurile pentru a lega materia de întrebări legate de valorile tehnologice, economice şi sociale. Dacă ar trebui să facă asta ar pierde un timp preţios. Cursurile lor ar putea deveni mai puţin atrăgătoare pentru studenţi. Universităţile vor să demonstreze că studenţii lor se pot angaja după absolvire, iar angajatorii vor absolvenţi cu aptitudini puternice. Cei Nouă Titani au parteneriate cu aceste universităţi, care depind de finanţarea şi resursele lor. Cu toate acestea, a venit momentul ca întrebările dificile – *cine e proprietarul tău*? – să fie puse şi dezbătute în mediul universitar, înainte ca studenţii să devină membri ai unor echipe aflate în mod constant sub presiunea termenelor limită şi a ţintelor de încasări.

Dacă universităţile sunt locul unde se nasc triburile inteligenţei artificiale, e uşor de văzut de ce în acest domeniu diversitatea este mult mai scăzută faţă de alte profesii. Directorii companiilor de tehnologii se pripesc să arate cu degetul înspre universităţi, dând vina pe ele pentru diversitatea scăzută a forţei de muncă. Nu e o acuzaţie complet falsă. Triburile inteligenţei artificiale se formează în laboratoare şi săli de clasă, în timp ce studenţii colaborează la diferite teme şi proiecte. Aceşti profesori, laboratoarele şi conducerea departamentelor de inteligenţă artificială din facultăţi sunt în proporţie covârşitoare compuse din bărbaţi şi cu o diversitate scăzută.

În cadrul universităţilor, doctoranzii îndeplinesc trei

funcţii: să colaboreze în activitatea de cercetare, să predea studenţilor şi să ducă mai departe domeniul. Femeile beneficiază de doar 23% din doctoratele în informatică şi doar 28% din doctoratele în matematică şi statistică, conform datelor recente ale Naţional Center for Education Statistics.74 Femeile dispar pe parcursul academic: doctorandele nu sunt promovate în roluri de conducere sau ca profesori titulari în egală măsură cu bărbaţii. Nu ar trebui deci să surprindă pe nimeni că femeile au reprezentat doar 18% din absolvenţii de informatică din ultimii ani – procentaj chiar mai scăzut decât cel din 1985, de 37%.75 Candidaţii negri sau latino-americani sunt şi ei foarte absenţi – constituind doar 3%, respectiv 1% din totalul doctoranzilor.76

Pe măsură ce tribul creşte, expansiunea se produce în interiorul unei bule, provocând unele comportamente inacceptabile. Cercetătoarele din domeniul inteligenţei artificiale au trebuit să se confrunte cu hărţuire sexuală, glume nepotrivite şi un comportament în general execrabil al colegilor lor. Pe măsură ce acest comportament devine considerat normal, el se transferă odată cu tribul dinspre universitate înspre forţa de muncă. Prin urmare, nu e atât o problemă a parcursului de instruire, cât o problemă de *oameni*. Triburile inteligenţei artificiale sprijină pur şi simplu o cultură din care femeile şi anumite minorităţi-negrii sau latino-americanii – sunt excluşi.

Un inginer Google a trimis o notă de informare devenită celebră în 2017 prin care susţinea că femeile sunt biologic inferioare când vine vorba de programare. Directorul general al Google, Sundar Pichai, a reacţionat în cele din urmă concediindu-l pe angajat, însă a afirmat şi că „putem dezbate o bună parte din acea notă de informare”. 77 Culturile ostile faţă de membrii din afara tribului amplifică acest efect, prin urmare forţa de muncă devine şi mai puţin diversă. Pe măsură ce inteligenţa artificială progresează, construind sisteme care gândesc pentru şi alături de oameni, segmente întregi din populaţie

sunt excluse.

Nu vreau să spun că în universităţi nu lucrează şi femei sau oameni de culoare. Directorul faimosului Computer Science and Artificial Intelligence Lab de la MIT este Daniela Rus, o femeie care a fost bursieră MacArthur, printre multe alte realizări profesionale şi academice. Kate Crawford este profesor emerit la New York University şi conduce un institut nou înfiinţat care studiază efectele inteligenţei artificiale asupra societăţii. Există şi femei şi oameni de culoare care fac o treabă grozavă în domeniul inteligenţei artificiale – numărul lor este însă foarte redus.

Dacă tribul urmăreşte să ofere inteligenţei artificiale o gândire mai „umanistă”, mulţi oameni sunt însă excluşi din acest proces. Fei-Fei Li, care conduce laboratorul de inteligenţă artificială de la Stanford şi este cercetător principal în domeniul inteligenţei artificiale şi învăţării maşinilor la Google Cloud spune că

„Ca femeie, ca educator, ca femeie de culoare, ca mamă, sunt din ce în ce mai îngrijorată. Inteligenţa artificială urmează să aducă cele mai mari schimbări din istoria omenirii, iar nouă ne lipseşte o generaţie întreagă de informaticieni şi lideri cu diversitate… Dacă nu reuşim să integrăm femeile şi persoanele de culoare – în rândul informaticienilor care lucrează pe bunevom prejudicia sistemele. Dacă vom încerca peste zece sau douăzeci de ani să revenim asupra acestei situaţii va fi mult mai dificil, dacă nu chiar aproape imposibil”. 78

Triburile Chinei: BAT

Baidu, Alibaba şi Tencent, cunoscute împreună sub numele de BAT reprezintă partea chineză a celor Nouă Titani. Tribul inteligenţei artificiale din Republica Populară Chineză operează cu reguli şi ritualuri diferite, pe baza unor finanţări, controale şi politici industriale ale guvernului care sunt menite să-i împingă pe membrii BAT spre viitor. Această structură bine finanţată şi organizată implementează planurile la nivel naţional pentru implementarea viitoare a inteligenţei artificiale, iar

guvernul deţine o putere extraordinară în cadrul structurii. Această epocă este echivalentul cursei spaţiale pentru China, iar noi suntem pentru ei ce reprezenta Sputnik pentru misiunea Apollo. Poate că am fost primii care am ajuns pe orbită, însă China a pus la bătaie fondul suveran de investiţii, sistemul educaţional, cetăţenii şi mândria naţională în cursa pentru inteligenţa artificială.

Şi triburile inteligenţei artificiale din China îşi au rădăcinile în sistemul universitar, unde accentul pe aptitudini şi aplicaţiile comerciale este şi mai pronunţat. Deoarece China este interesată să sporească cât mai repede aptitudinile forţei de muncă, problemele de diversitate nu sunt tocmai comparabile cu cele ale Occidentului, deşi ele există. Genul nu joacă un rol atât de mare, astfel încât femeile sunt mai bine reprezentate. Cu toate acestea, cursurile sunt predate în chineză, o limbă dificil de stăpânit pentru străini. Se exclud astfel persoanele nevorbitoare de chineză, creându-se însă şi un avantaj concurenţial special, deoarece studenţii chinezi cunosc în general limba engleză şi pot studia în mai multe universităţi.

În China, instruirea în inteligenţă artificială începe înainte ca studenţii să intre la facultate. În 2017, consiliul de stat al Chinei a cerut includerea unor cunoştinţe de bază în inteligenţa artificială în programa şcolară, ceea ce înseamnă că elevii chinezi încep să se pregătească din timpul *şcolii primarei*. Există în prezent un manual oficial editat sub egida guvernului care prezintă istoria şi elementele de bază ale inteligenţei artificiale. Până în 2018, 40% dintre licee au luat parte la un program pilot cu un curs obligatoriu de inteligenţă artificială, 79 şi odată ce vor fi disponibili mai mulţi profesori şi alte şcoli vor fi incluse în program. Nu va dura prea mult pentru că ministerul educaţiei din China a lansat un program de instruire în inteligenţă artificială în universităţi, cu peste 500 de profesori şi 5.000 de studenţi de la universităţile de vârf din China.80

Membrii BAT contribuie la revoluţia educaţiei din China, oferind instrumentele folosite în şcoli şi universităţi, vânzând pe piaţă produsele pe care consumatorii le folosesc în copilărie şi la maturitate, angajând absolvenţii şi făcând schimb de experienţă cu guvernul. Dacă nu aţi călătorit sau trăit în China în ultimii 10 ani, poate că Baidu, Alibaba şi Tencent nu sunt nişte nume prea cunoscute. Toate cele trei companii au fost înfiinţate în aceeaşi perioadă, folosind ca model companiile din domeniul tehnologiei care existau în acea perioadă.

Baidu a fost înfiinţată în cadrul unui picnic de vară din Silicon Valley-una dintre acele ocazii exclusiviste în care membrii tribului inteligenţei artificiale se adună şi beau bere. Trei bărbaţi, toţi pe la vreo 30 de ani, se plângeau de cât de puţin evoluează motoarele de căutare. John Wu, la vremea respectivă directorul echipei de motoare de căutare a Yahoo şi Robin Li, inginer la Infoseek, credeau că motoarele de căutare au înainte un viitor strălucit. Văzuseră deja ce rezultate avea o firmă nou înfiinţată-Google – şi credeau că pot face ceva asemănător în China. împreună cu biochimistul Eric Xu, cei trei au înfiinţat compania Baidu.81

Compania a recrutat angajaţi din centrele universitare de cercetare în inteligenţă artificială din America de nord şi China. Compania se pricepea îndeosebi să atragă cercetători cu aptitudini în domeniul învăţării profunde. În anul 2012, compania i-a făcut o ofertă lui Andrew Ng, un cercetător de vază de la divizia Brain a lui Google. el a crescut în Hong Kong şi Singapore şi a trecut prin centrele importante în inteligenţă artificială: diplomă de licenţă la Carnegie Mellon, diplomă de master la MIT, doctorat la UC Berkeley, apoi o carieră didactică la Stanford. Ng le-a atras atenţia celor de la Baidu pentru că lucra la un proiect revoluţionar al unei reţele neurale profunde la Google.

Echipa lui Ng construise o reţea de 1.000 de calculatoare care s-a învăţat singură să recunoască pisicile

în videoclipurile de pe Youtube. Era un sistem remarcabil. Fără să i se fi spus niciodată în mod explicit ce este o pisică, inteligenţa artificială a vizionat milioane de ore de videoclipuri la întâmplare, a învăţat să recunoască obiecte, şi-a dat seama că anumite obiecte sunt pisici apoi a învăţat ce este o pisică. Toate pe cont propriu, fără ca oamenii să aibă vreo influenţă. La puţin timp după aceea, Ng s-a angajat la Baidu ca cercetător principal. (Acest lucru înseamnă că în ADN-ul Baidu se regăsesc obligatoriu nucleotide din cursurile de inteligenţă artificială de la Carnegie Mellon, MIT şi UC Berkeley.)

În ziua de azi, Baidu nu e doar un motor de căutare. Ng a contribuit la platforma de conversaţii bazată pe inteligenţă artificială (Ducros), un asistent digital şi programe autonome de condus maşina, precum şi alte aplicaţii ale inteligenţei artificiale care au permis Baidu să înregistreze progrese în domeniu cu mult înaintea Google. Baidu are o capitalizare de piaţă de 88 de miliarde de dolari şi e al doilea motor de căutare din lume după Google – o mare realizare având în vedere că nu este folosit în afara Chinei. La fel ca Google, Baidu dezvoltă o gamă de aparate casnice inteligente, cum ar fi un robot domestic care îmbină recunoaşterea vocală cu cea facială. Compania a anunţat o platformă deschisă pentru conducerea autonomă a vehiculelor intitulată Apollo, având speranţa ca statutul deschis al codului sursă să provoace o înflorire a ecosistemului limitrof. Baidu colaborează cu peste 100 de parteneri, printre care se numără constructorii de automobile Ford şi Daimler, fabricanţii de chipuri NVIDIA şi Intel şi firme de cartografie precum Tomâom. Baidu a încheiat un parteneriat cu compania californiană Access Services pentru a pune la dispoziţia persoanelor cu probleme de mobilitate maşini autonome. Baidu a încheiat un parteneriat şi cu divizia Azure Cloud a Microsoft pentru a permite partenerilor Apollo din afara Chinei să proceseze cantităţi uriaşe de date despre maşini.82 Ar trebui să ştiţi şi că Baidu a inaugurat în ultimii ani un laborator nou de

cercetare în inteligenţă artificială în parteneriat cu guvernul chinez – iar liderii laboratorului sunt nomenclaturişti cu experienţă în programele armatei.83

A-ul din BAT corespunde Alibaba Group, o platformă uriaşă care joacă rol de intermediar între cumpărători şi vânzători printr-o reţea gigantică de site-uri, nu printr-o platformă unică. Alibaba a fost înfiinţat în 1999 de către Jack Ma, un fost profesor care trăieşte la aproximativ 150 de km la sud-vest de Shanghai, care şi-a dorit să creeze o variantă hibrid între Amazon şi ebay pentru piaţa din China. Ma nu avea cunoştinţe informatice, aşa că a înfiinţat compania împreună cu un coleg de facultate. După doar 20 de ani, Alibaba are o capitalizare de piaţă de peste 511 miliarde de dolari.

Printre site-urile deţinute de Alibaba se numără Taobao, pe care atât cumpărătorii cât şi vânzătorii nu plătesc nicio taxă pentru tranzacţiile efectuate. Taobao foloseşte în schimb un model, prin care taxează vânzătorii care doresc să obţină un loc mai bun pe motorul de căutare al site-ului. (Această tactică imită o parte din modelul de afaceri al Google.) Alibaba a dezvoltat şi sisteme de plată sigure, printre care se numără Alipay, cu o funcţionalitate şi caracteristici asemănătoare cu PayPal. Alibaba a lansat un sistem de plăţi digitale „zâmbeşte ca să plăteşti” bazat pe inteligenţă artificială şi recunoaştere facială, care permite consumatorului să plătească zâmbind în faţa unei camere.

La fel ca Amazon, Alibaba are un difuzor inteligent – se numeşte Genie XI, e mai mic şi îndesat de Alexa de la Amazon sau aparatele Home ale Google. Genie XI foloseşte o tehnologie de recunoaştere a amprentei vocale bazată pe reţele neurale pentru a identifica utilizatorii, autentificându-i automat pentru a le permite să vizualizeze produse şi să le cumpere. Peste 100.000 de astfel de aparate Alibaba sunt în curs de instalare în hotelurile Marriott din toată China.

Alibaba are o viziune de ansamblu asupra inteligenţei

artificiale, pe care o denumeşte „ET City Brain”. Programul prelucrează cantităţi uriaşe de date obţinute local, de la camere şi senzori inteligenţi din oraş, la arhive guvernamentale şi conturi de pe reţele de socializare. Alibaba foloseşte această infrastructură de inteligenţă artificială pentru a genera modele predictive: pentru a îmbunătăţi traficul, urbanismul, sănătatea publică şi a stabili dacă urmează să se producă tulburări sociale. Sub conducerea lui Ma, Alibaba a intrat şi în domeniul logisticii, găzduirii de videoclipuri online, al centrelor de stocare a datelor şi tehnologiilor cloud, investind miliarde de dolari în diferite companii, urmărind să construiască un colos digital care să conecteze lumea de afaceri, casele, munca, oraşele şi guvernul. De fapt, înainte ca magazinul Amazon Go să fie inaugurat în Seattle, Alibaba lansase magazinul Hema, un concept automat, fără bani gheaţă, care combină un magazin alimentar, o piaţă şi un serviciu de livrare la domiciliu.

Mai există o asemănare stranie care merită amintită aici. Spun „stranie” pentru că ea reprezintă şi o contradicţie. În anul 2016, Ma a cumpărat ziarul *South China Morning Post*, cel mai mare şi mai influent cotidian din Hong Kong. Tranzacţia a fost una importantă pentru că în China majoritatea publicaţiilor sunt finanţate de stat iar acest ziar de limbă engleză a fost cunoscut pentru articolele critice la adresa guvernului chinez.84 Când trăiam în Hong Kong, obişnuiam să beau cu mai mulţi reporteri de la acest ziar, nişte jurnalişti de investigaţii de primă mână. Prin achiziţia acestui ziar, Ma îşi demonstra loialitatea faţă de partidul comunist. Cu trei ani în urmă, Jeff Bezos cumpărase *Washington Post*, o mutare care l-a transformat în cele din urmă într-un inamic al Casei Albe conduse de preşedintele Trump din cauza jurnalismului obiectiv al ziarului, analizei critice asupra politicilor acestei administraţii şi a demersurilor continue de combatere a propa85

În ultimul rând, cel mai mare şi în multe privinţe cel

mai influent membru al BAT este Tencent. Această companie a fost înfiinţată în 1998 de doi bărbaţi, Ma Huateng şi Zhang Zhidong. Aceştia au pornit de la un singur produs denumit OICQ. Dacă vi se pare că sună uşor cunoscut, asta e pentru că era o copie a serviciului de mesagerie ICQ. Cei doi au fost daţi în judecată dar au continuat să lucreze la propria variantă a sistemului. În 2011, Tencent a lansat WeChat, care oferă nu doar un serviciu de mesagerie ci copiază şi funcţionalităţile şi caracteristicile Facebook. Din momentul în care guvernul chinez a luat decizia de a bloca Facebook în China, WeChat a dobândit o poziţie avantajoasă. Platforma a depăşit graniţele mediului academic, fiind folosită pentru recrutarea noilor angajaţi-şi în multe alte scopuri.

WeChat are o cifră uluitoare de 1 miliard de utilizatori activi în fiecare lună şi o poreclă „aplicaţia pentru orice”. Asta pentru că dincolo de postările standardizate pe reţelele de socializare şi mesagerie, WeChat este folosită pentru aproape orice în China, de la recrutarea noilor angajaţi în universităţi şi mesagerie text la plăţi virtuale şi chiar măsuri de ordine publică. Peste 38.000 de spitale şi clinici au conturi pe WeChat, şi 60% dintre acestea le folosesc pentru gestionarea relaţiei cu pacienţii (de ex., programări şi plăţi).86 Tencent este o companie care se sprijină – şi se concentrează – pe inteligenţa artificială, considerând că ea „reprezintă o tehnologie fundamentală în toate produsele noastre diferite”. 87 Nu e de mirare că sloganul oficial al companiei este „Inteligenţa artificială peste tot”.

Facebook ar putea fi cea mai mare reţea de socializare din lume, însă tehnologia de care dispune Tencent este din multe puncte de vedere superioară. Tencent a construit un asistent digital numit Xiaowei, un sistem mobil de plăţi (TenPay), un serviciu cloud (Weiyun) şi a lansat recent şi un studio de film (Tencent Pictures). Laboratorul Youâu al Tencent obţine rezultate de vârf în recunoaşterea facială şi a obiectelor, şi transferând aceste

tehnologii către peste 50 de alte departamente ale companiei. Tencent are şi iniţiative în domeniul sănătăţii, încheind parteneriate cu două companii de sănătate din Marea Britanie: Babylon Health, o firmă nou înfiinţată din domeniul medicinei la distanţă şi Medopad, care foloseşte inteligenţa artificială pentru monitorizarea pacienţilor de la distanţă. Tencent a investit masiv şi în două companii promiţătoare din Statele Unite, Atomwise şi Xtal Pi care vizează aplicarea inteligenţei artificiale în domeniul farmaceutic.

Tencent a devenit prima companie din Asia care a depăşit în 2018 o valoare de piaţă de 550 de miliarde de dolari şi a depăşit Facebook, devenind cea mai valoroasă reţea de socializare din lume.88 Ceea ce e şi mai uimitor e că sub 20% din veniturile Tencent provin din publicitatea online, în vreme ce Facebook depinde în proporţie de 98% de aceste venituri.89

Pepiniera BAT include centrele universitare din America de nord, iar compania se asigură că copiii primesc o educaţie în inteligenţă artificială începând cam de la vârsta la care învaţă să adune şi să scadă.

Toate aceste aspecte nu ar fi atât de importante dacă grupul BAT nu s-ar bucura de un succes extraordinar şi dacă nu ar câştiga sume uriaşe de bani. BAT produce atât de mult iar piaţa din China este atât de mare, încât triburile inteligenţei artificiale din China au o putere uriaşă – atât în China cât şi în alte părţi ale lumii. Comunitatea globală a inteligenţei artificiale e atentă la China din cauza acestui capital şi a cifrelor care sunt greu de ignorat.

Facebook ar putea avea 2 miliarde de utilizatori activi în fiecare lună, dar ei sunt împrăştiaţi în toată lumea. Miliardul de utilizatori WeChat se găseşte în mare măsură într-o singură ţară. Baidu are 665 de milioane de utilizatori care folosesc căutarea mobilă în 201790 – mai mult decât dublul numărului estimat de deţinători de telefoane mobile în Statele Unite.91 în acelaşi an, Amazon

a înregistrat cele mai bune vânzări de sărbători din istorie. Pentru a putea înţelege, între ziua Recunoştinţei şi Cyber Monday, b clienţii Amazon au comandat 140 de milioane de produse, cu o valoare totală de 6, 59 miliarde de dolari.92 Poate că pentru Amazon ar fi fost un record, însă nici nu se compară cu realizările Alibaba în China, unde în doar 25 de ore, Alibaba a vândut unui număr de 515 milioane de clienţi în 2017, iar de ziua oamenilor singuri, un soi de Black Friday combinat cu premiile Oscar în China – a vândut de 25 de miliarde de dolari pe internet, procesând 812 milioane de comenzi *într-o singură zi.93* După orice standarde, China are cea mai mare piaţă digitală din lume: peste 1.000 de miliarde de dolari cheltuiţi anual, peste un miliard de oameni pe internet, şi 30 de miliarde de dolari investite speculativ în cele mai importante companii din domeniul tehnologic din lume.94

Investitorii chinezi au fost implicaţi în 7% până la 10% din toate finanţările companiilor nou înfiinţate în domeniul tehnologic din Statele Unite în perioada 2012 – 2017 – vorbim de sume semb Ziua Recunoştinţei se sărbătoreşte în Statele Unite în ultima zi de joi a lunii noiembrie, iar Cyber Monday este prima zi de luni de după Ziua Recunoştinţei (n.t.) nificative de bani care provin dintr-o singură regiune.95 BAT au acum o prezenţă palpabilă în Seattle şi Silicon Valley, cu propriile filiale, inclusiv birouri pe faimoasa stradă Sand Hill Road. În ultimii cinci ani, BAT a investit sume considerabile în companii ca Tesla, Über, Lyft, Magic Leap (un fabricant de căşti de realitate virtuală) şi altele. Investiţiile companiilor din grupul BAT sunt atrăgătoare nu doar pentru că se mişcă rapid şi au mulţi banici şi pentru că un acord cu grupul BAT presupune de regulă o portiţă profitabilă către piaţa din China, unde altminteri ar fi imposibil de pătruns. Spre exemplu Zoloz, o companie mică din Kansas City care se ocupa cu recunoaşterea facială, a fost cumpărată de Alibaba pentru suma de 100 de milioane de dolari în 2016, devenind apoi o componentă de bază a sistemului de plăţi

Alipay, obţinând acces la sute de milioane de utilizatori fără a trebui să se confrunte cu legile draconice legate de intimitate din Europa, sau cu ameninţarea unor potenţiale procese în Statele Unite. Investiţia nu a venit însă lipsită de compromisuri serioase. Investitorii chinezi nu aşteaptă doar ca investiţia să dea randamentei cer şi drepturile de proprietate intelectuală asupra tehnologiei.

În China, faptul de a pretinde să primeşti drepturile de proprietate intelectuală asupra tehnologiei în schimbul banilor nu reprezintă o particularitate culturală, sau o modalitate lacomă prin care anumiţi investitori să câştige mai mult. Această abordare reprezintă un demers coordonat al guvernului. China are o viziune clară asupra felului în care vrea să domine global din punct de vedere economic, geopolitic şi militarşi consideră că inteligenţa artificială este calea către acest obiectiv. În vederea acestui scop, menţinerea unui control autoritar asupra informaţiilor reprezintă o problemă foarte importantă pentru conducătorii ţării, aşa că China a adoptat o politică industrială menită să transfere proprietatea intelectuală de la companiile din Statele Unite către cele chineze. Printre exemple se numără anumite seturi de date, algoritmi sau designul procesoarelor. Multe companii americane care speră să facă afaceri în China trebuie să promită mai întâi să ofere acces la propriile tehnologii. Au fost adoptate noi reglementări, prin care companiile străine trebuie să-şi desfăşoare activitatea de cercetare şi dezvoltare în China şi să stocheze datele folosite local. Stocarea locală a datelor este o cerinţă dură impusă companiilor din străinătate deoarece guvernul chinez ar putea să invoce autoritatea sa pentru a accesa oricând aceste date, ocolind metodele de criptare.

Beijingul ia în serios planificarea pe termen lung. Tradiţia a început încă din vremea lui Mao Zedung, care a introdus primele cincinaluri începând cu anul 1953. (Preşedintele Xi a lansat cel de-al 13-lea cincinal în 2016).96 Atât liderii politici cât şi nomenclaturiştii sunt

adepţii prevederii strategicefăcând din China una dintre puţinele ţări de pe glob care are o strategie economică, politică, militară şi socială amănunţită pentru următoarele decenii. Guvernul chinez are capacitatea unică de a implementa orice politică îşi doreşte şi de a face orice este necesar pentru a-şi pune în aplicare strategia naţională, inclusiv planul de a transforma până în 2030 China în „principalul centru de inovare în inteligenţa artificială din lume” şi să creeze o industrie în valoare de 150 de miliarde de dolari. E puţin probabil ca acest plan să fie anulat de către un nou guvern, deoarece China a abrogat în luna martie 2018 limitele privind mandatele, permiţând preşedintelui Xi Jinping să rămână la putere pe viaţă.

Sub Xi, China a înregistrat o consolidare impresionantă a puterii. El a întărit Partidul Comunist, a suprimat fluxul informaţiilor şi a instituit noi politici pentru accelerarea unei multitudini de planuri pe termen lung, de pe urma cărora se aşteaptă să beneficieze în deceniul următor. La nivelul cel mai de sus al guvernului chinez, inteligenţa artificială ocupă o poziţie centrală. Spre deosebire de fostul lider al Partidului Comunist Chinez, Deng Xiaoping, care guverna în baza principiului „să ne ascundem capacităţile şi să tragem de timp”, Xi e gata să arate lumii ce poate China să facă – şi intenţionează să preia iniţiativa la nivel global.97 Conducerea Chinei e orientată spre viitor şi are de gând să-şi pună în aplicare chiar acum planurile îndrăzneţe, unite. Acest aspect oferă Chinei un avantaj incredibil faţă de Occident, iar membrilor grupului BAT le conferă o putere extraordinară.

Toate acestea se petrec într-o perioadă de creştere economică puternică a Chinei, care înregistrează o creştere extraordinară a clasei de mijloc. Până în 2022, mai bine de trei sferturi din populaţia urbană a Chinei va dispune de suficiente resurse pentru a face parte din clasa de mijloc. În 2000, doar 4% erau consideraţi a face parte din clasa de mijlocvorbim de o creştere uriaşă într-un interval scurt de timp. Slujbele mai bine plătite din

domeniul informaticii, biotehnologiilor şi serviciilor vor ridica o mare parte a acestui grup din pătura de mijloc în partea superioară a clasei de mijloc. Gospodăriile din China au un grad de îndatorare foarte redus. Deşi e adevărat că peste tot în China există şi sărăcie, generaţia actuală a copiilor din China are oportunităţi bune de a câştiga mai mult, a economisi mai mult şi a cheltui mai mult decât părinţii lor.98 (în mod frapant, 70% dintre americani se consideră parte a clasei de mijloc, dar datele Pew Research Center" arată că clasa de mijloc americană s-a micşorat în ultimele patru decenii – mai puţin de jumătate dintre americani câştigă suficient pentru a se putea încadra în această categorie.100)

China a devenit o putere economică greu de ignorat. Compania Marriott a semnat un contract pentru instalarea a 100.000 de aparate inteligente Alibaba în hotelurile din China, însă când Beijingul a aflat că pe site-ul companiei Hong Kong, Taiwan, Tibet şi Macau sunt trecute ca ţări diferite într-un chestionar trimis pe e-mail membrilor clubului de clienţi fideli, directorii Marriott au primit o somaţie imediată. Guvernul i-a comunicat companiei Marriott să-şi închidă toate site-urile şi aplicaţiile din China, iar compania s-a supus. Compania are un portofoliu de peste 240 de hoteluri şi staţiuni de lux deschise în toată China pentru a profita de expansiunea clasei de mijloc. Directorul general Arne Sorenson a fost nevoit să publice un comunicat uluitor pe site-ul companiei: „Marriott International respectă şi sprijină suveranitatea şi integritatea teritorială a Chinei. Din nefericire, în ultima săptămână au avut loc două incidente care nu confirmă această abordare: mai întâi, printr-o etichetare incorectă a unor regiuni din China, inclusiv a Tibetului, ca fiind ţări într-un meniu dintr-un sondaj pe care-l trimitem clienţilor fideli, iar apoi prin faptul că unul dintre angajaţii noştri a apreciat o postare care sugera în mod incorect că am sprijini acest punct de vedere. Nimic nu e mai departe de adevăr: nu sprijinim pe nimeni care subminează

suveranitatea şi integritatea teritorială a Chinei şi nu dorim în vreun fel să încurajăm sau să incităm aceşti oameni sau aceste grupuri. Recunoaştem gravitatea situaţiei şi ne cerem sincer scuze”. 101

China a devenit şi o forţă geopolitică prea puternică pentru a fi subminată. Ea pune presiune asupra guvernelor străine printr-un plan de politică externă menit să revitalizeze pentru secolul 21 fostul drum multimilenar al mătăsii. China cheltuie 150 de miliarde de dolari pe an în 68 de ţări pentru repararea infrastructurii: drumuri, căi ferate, poduri, porturi. Prin aceste programe ţările participante se vor înscrie din ce în ce mai mult pe orbita Beijingului, într-o perioadă în care America bate în retragere. În timp America e ocupată cu incertitudinile şi haosul administraţiei Trump, preşedintele Xi a făcut din China un pol al stabilităţii. În absenţa Americii, Xi a început să umple golul de putere la nivel global.

Donald Trump, de pildă, a repetat pe Twitter în timpul campaniei afirmaţii care neagă încălzirea globală, inclusiv referitoare la o teorie bizară a conspiraţiei conform căreia chinezii ar dori să ne saboteze economia.102 Evident că nu există niciun sâmbure de adevăr în toată povestea asta. China a construit alianţe în ultimul deceniu, urmărind să reducă deşeurile de plastic, să obţină mai multă energie regenerabilă şi să elimine poluarea industrială. China nu prea avea de ales: după decenii în care a fost inima industrială şi groapa de gunoi a lumii, poluarea a provocat probleme de sănătate generalizate şi reducerea speranţei de viaţă în China, în 2017, guvernul a anunţat că China nu va mai importa deşeuri din alte ţări, după ce, începând cu 1992, China a importat şi procesat 106 milioane tone de deşeuri.103 Fiindcă Statele Unite nu îşi fac planuri pe termen lung, nu avem la dispoziţie o alternativă. În prezent nu avem la dispoziţie alternative pentru depozitarea deşeurilor, ceea ce înseamnă că China forţează celelalte ţări să nu mai folosească bunuri care nu pot fi reciclate. China devine

rapid un lider global în problema dezvoltării durabile, fiind suficient de puternică pentru a-şi impune punctul de vedere.

În China, oamenii apreciază *chengyu*, nişte aforisme din patru ideograme prin care transmit înţelepciunea. Unul care-mi vine acum în minte este Wiffii tu, care în traducere înseamnă „bobul se leapădă de pleavă şi dă rod”. 104 China îşi arată acum lumii foarte făţiş puterea de o manieră foarte evidentă.

Consolidarea puterii lui Xi, combinată cu puterea şi ascensiunea economică a Chinei pun bazele înfloririi triburilor inteligenţei artificiale, în special dat fiind demersul consolidat vertical în domeniul inteligenţei artificiale. La marginea Beijingului se construieşte un parc de cercetare în valoare de 2 miliarde de dolari care se va concentra pe învăţare profundă, tehnologii cloud şi biometrice şi va include un laborator naţional de cercetare şi dezvoltare. Nu doar că guvernul investeşte în grupul BAT ci îl şi protejează de concurenţa globală. Guvernul chinez interzice accesul Google şi Facebook, şi a zădărnicit încercarea Amazon de a intra pe piaţă. Companiile din grupul BAT se află în centrul planului guvernamental de dezvoltare până în 2030, care se bazează puternic pe tehnologiile acestor companii: sistemele autonome de condus ale Baidu, tehnologia Internet of Things a Alibaba şi sistemele comerciale interconectate şi aplicaţiile Tencent în domeniul interfeţelor de comunicare şi asistenţei sanitare.

Iată de ce ar trebui să vă preocupe triburile inteligenţei artificiale din China, indiferent în ce ţară locuiţi.

Economia Chinei creşte într-un ritm alert, iar evoluţia rapidă a inteligenţei artificiale nu face decât să accelereze ascensiunea Chinei. La finalul anului 2017, analiza şi modelele pe care le-am efectuat împreună cu echipa de la Future Today Institute au arătat că inteligenţa artificială are potenţialul de a spori economia Chinei cu 28% până în

2035. Inteligenţa artificială – alimentată de numărul uriaş de cetăţeni chinezi şi datele lor, automatizarea la scară largă, maşinile care învaţă şi eficientizarea cheltuielilor de capital – va stimula creşterea industriei, agriculturii, comerţului, serviciilor financiare, transporturilor, utilităţilor, sănătăţii şi divertismentului (inclusiv a platformelor) din China. În momentul de faţă nu există o altă ţară cu atât de multe date, atât de mulţi oameni şi atât de multe aparate electronice pe cap de locuitor. Nicio altă ţară nu are şansa de a depăşi Statele Unite în decurs de o generaţie. Nicio altă ţară nu are un potenţial mai mare de a influenţa ecosistemul şi clima planetei – conducând fie la supravieţuire, fie la catastrofadecât China. Nicio altă ţară nu constituie o punte între lumea dezvoltată şi cea în curs de dezvoltare în aceeaşi măsură ca China. Cu un regim comunist dar o economie puternică, China e atât un partener prea mare pentru a putea fi ignorat, un adversar politic cu puncte de vedere complet diferite în ce priveşte drepturile omului, dar şi un coagulant al alianţelor globale. Prosperitatea economică aduce puterea. China se poziţionează pentru a putea influenţa fluxurile financiare şi comerciale mondiale. Prin această reaşezare celelalte ţări îşi pierd puterea şi influenţa, iar democraţia are de suferit în lume.

În al doilea rând, China se va folosi de progresele din domeniul inteligenţei artificiale şi de stimulente economice pentru a-şi moderniza armata, oferindu-i un avantaj asupra ţărilor din Occident. Această tranziţie a început deja ca parte a unui program naţional de supraveghere aeriană cu numele de cod „Porumbelul”. Peste 30 de agenţii militare şi guvernamentale au lansat „păsări spion” care arată ca nişte păsări albe care bat imitând aripile păsărilor. Dronele fac parte dintr-un program inspirat din natură menit să evite radarul şi detecţia.105 Dronele filmează, iar un sistem de inteligenţă artificială caută tipare, recunoaşte feţe şi identifică anomalii. Cu toate acestea, deşi păsările spion sună ameninţător, ele sunt cel mai mic motiv de

îngrijorare pe care l-aţi putea avea.

La finalul anului 2017, un raport nepublicat al Pentagonului obţinut de către reporterii Reuters avertiza că firmele chineze se sustrag supravegherii guvernamentale a Statelor Unite şi accesează tehnologii secrete din domeniul inteligenţei artificiale din Statele Unite care ar putea avea şi aplicaţii militare cumpărând companii din Statele Unite. Armata Populară de Eliberare a Chinei investeşte puternic într-o gamă de proiecte şi tehnologii din domeniul militar, în timp ce institutele de cercetare militare colaborează cu industria de apărare din China.106

China nu a mai luptat împotriva unei alte ţări din 1979, cu ocazia războiului chino-vietnamez. S-ar părea că China nu are adversari militari de calibru – nu a fost ţinta unor atacuri teroriste, nu are relaţii antagonice cu suspecţii de serviciu (de ex., Rusia, Coreea de Nord) şi nu are alţi duşmani. De unde atunci această militarizare intensă?

Motivul este că, în viitor, războaiele vor fi purtate prin intermediul programelor, nu prin lupte la baionetă. Folosind tehnologiile inteligenţei artificiale, o armată poate „câştiga” destabilizând o economie, în loc să pustiască satele şi oraşele. Dată fiind această perspectivă şi demersurile puternice ale Chinei în domeniul inteligenţei artificiale, China are un avans periculos faţă de Occident.

Din punctul meu de vedere, ne-am dat seama prea târziu de acest lucru. Când m-am întâlnit cu oficialităţile de la Pentagon a fost dificil ca perspectiva nouă asupra războiului (programe versus luptă) să fie acceptată. De exemplu, în 2017, departamentul Apărării a înfiinţat o echipă interdisciplinară de război algoritmic pentru a contribui la proiectul Maven – un sistem de vizualizare digitală şi învăţare profundă care recunoaşte automat obiecte din imagini şi videoclipuri. Echipa nu dispunea de resursele necesare în inteligenţă artificială aşa că departamentul a încheiat un contract cu Google pentru a

pune la punct sistemele care să analizeze imaginile filmate din drone. Angajaţilor Google nu li s-a spus însă că vor lua parte la un proiect militar, iar reacţiile au fost vehemente. Patru mii de angajaţi Google au semnat o petiţie prin care îşi exprimau dezacordul faţă de proiectul Maven. Ei au cumpărat o pagină de publicitate în *New York Times*, şi în cele din urmă zeci de angajaţi au demisionat.107 Google a comunicat că nu va reînnoi contractul cu departamentul Apărării.

Şi Amazon a fost luată în vizor din cauza unui contract cu Pentagonul în valoare de 10 miliarde de dolari. În luna octombrie 2018, Tom Cole, republican, membru al comisiei pentru buget din Oklahoma, şi Steve Womack, republican din Arkansas, au acuzat Departamentul Apărării că a redactat caietul de sarcini în colaborare cu Amazon pentru a exclude alte companii. Aceasta nu a fost însă singura acuzaţie. Unii dintre angajaţii Amazon s-au scandalizat la aflarea veştii că angajatorul lor lucrează pentru armata Statelor Unite, în vreme ce alţii nu au apreciat faptul că tehnologiile de recunoaştere facială sunt puse la dispoziţia forţelor de ordine. Ca răspuns, Jeff Bezos a declarat în faţa unei conferinţe de presă: „Dacă companiile mari din domeniul tehnologiei vor întoarce spatele Departamentului Apărării, ţara noastră va avea mari probleme”. 108

Dacă în Statele Unite giganţii tehnologici încearcă să jongleze între transparenţă şi securitatea naţională, relaţia dintre grupul BAT şi guvernul chinez se află la polul opus. Iată însă un exemplu îngrijorător: poziţia actuală a armatei americane este că oamenii trebuie să deţină controlul, indiferent cât de avansate devin inteligenţa artificială, sistemele autonome şi roboţii. Prin acest control ne asigurăm că nu lăsăm decizia de a folosi forţa letală în mâinile unui program. China are o abordare diferită.109 General-locotenentul Liu Guozhi, şeful comisiei de tehnologie şi ştiinţă a armatei chineze ar fi avertizat” (noi) trebuie… să preluăm iniţiativa pentru a schimba

paradigma”. 110 Afirmaţia reprezintă o modalitate indirectă de a anunţa intenţia Chinei de a-şi reconstrui puterea militară.

În al treilea rând, dacă nu vă îngrijorează superioritatea economică şi militară a Chinei, perspectiva lor asupra intimităţii ar trebui să vă îngrijoreze. Din nou, de ce v-ar interesa atât timp cât nu locuiţi acolo? Pentru că guvernele autoritare se formează tot timpul, iar ele au tendinţa de a copia regimurile deja existente. Dată fiind creşterea la nivel global a naţionalismului, felul în care China se foloseşte de inteligenţa artificială ar putea deveni un model şi pentru alte state în anii care vor urma. Acest fapt ar putea dezechilibra pieţele, comerţul şi balanţa geopolitică de putere.

China se foloseşte de inteligenţa artificială în încercarea de a crea o populaţie mai obedientă, unul dintre cele mai vaste şi mai insidioase experimente din istoria omenirii. Planul consiliului de stat pentru inteligenţă artificială până în 2030 explică faptul că inteligenţa artificială va „îmbunătăţi semnificativ eficienţa şi nivelul guvernării sociale” şi va fi folosită pentru a juca „un rol de înlocuit în menţinerea eficientă a stabilităţii sociale”. 111 Strategia se implementează prin intermediul sistemului chinezesc naţional de credite sociale, care, conform cartei de înfiinţare a consiliului de stat, va „permite celor de încredere să se plimbe peste tot sub cer, iar cei discreditaţi abia vor face un singur pas”. 112 Această idee a apărut încă din 1949, când Partidul Comunist a preluat puterea şi a început să experimenteze diferite modalităţi de control social, în vremea regimului lui Mao Zedung din anii ’50, controlul strict al societăţii a devenit norma: muncitorii au fost colectivizaţi şi primeau calificative în funcţie de productivitate. Oamenii se urmăreau unii pe alţii iar ierarhia calificativelor stabilea cât de mult puteau accesa bunurile publice. Sistemul s-a descompus în vremea lui Mao, apoi din nou în anii ’80 pentru că, după cum s-a dovedit, oamenii nu se pricep de minune să se evalueze

reciproc, fiind motivaţi de propriile nevoi, nesiguranţe şi prejudecăţi.

În 1995, preşedintele Jiang Zemin şi-a imaginat un sistem de coerciţie socială bazat pe tehnologie – iar după 2005, guvernul chinez a lucrat pentru a construi şi implementa un sistem de notare care funcţiona automat.113 în parteneriat cu Universitatea din Beijing, guvernul a înfiinţat centrul de cercetare pentru credite din China, pentru a cerceta modalităţile de construire şi implementare a unui sistem naţional de credite sociale bazat pe inteligenţa artificială. Iată deci o explicaţie parţială a insistenţelor preşedintelui actual Xi pentru dezvoltarea inteligenţei artificiale. Prin dezvoltarea acestui sistem nu s-ar împlini doar un ideal de la înfiinţarea Partidului Comunist, ci implementarea lui ar reprezenta chiar un element care ar sprijini menţinerea partidului la putere.

Un sistem algoritmic de credite sociale a dovedit deja că inteligenţa artificială funcţionează în oraşul Rongeheng. Cei 740.000 de locuitori adulţi ai oraşului primesc la înscrierea în sistem 1.000 de puncte, iar funcţie de comportamentul demonstrat punctajul poate fi scăzut sau crescut. O faptă „eroică” poate aduce unui locuitor 30 de puncte, în timp ce a trece pe culoarea roşie a semaforului atrage o penalizare de 5 puncte. Cetăţenii sunt etichetaţi şi împărţiţi în diferite categorii, de la A ++ la D, iar alegerile pe care le pot face şi capacitatea de a se deplasa sunt dictate de categoria din care fac parte. Cei care fac parte din categoria C ar putea constata că li se solicită o garanţie dacă vor să închirieze o bicicletă, în timp ce cei din categoria A le pot închiria pe gratis timp de 90 de minute. Sistemul nu se aplică doar persoanelor fizice. În Rongeheng, şi *companiile* primesc note în funcţie de comportamentiar capacitatea lor de a opera depinde foarte mult de categoria din care fac parte.114

Străzile din Shanghai sunt împânzite de microfoane şi camere inteligente conectate la un sistem bazat pe

inteligenţă artificială. Şoferii care claxonează excesiv sunt amendaţi automat şi primesc amenda prin serviciul WeChat al Tentent, în timp ce numele, fotografia şi cartea de identitate sunt afişate pe panouri luminoase aflate în apropiere. Dacă un şofer îşi lasă maşina parcată mai mult de şapte minute, el primeşte automat amendă.115 Nu e vorba doar de amendă-şoferul este penalizat şi în sistemul de credite sociale. După ce se înregistrează abateri repetate, persoana ar putea întâmpina probleme în a-şi cumpăra bilete de avion sau a se angaja. Un episod popular al serialului *Black Mirror* ilustrează un viitor distopic similar. În Shanghai acest viitor a devenit deja realitate.

Supravegherea la nivel naţional se efectuează pe baza tehnologiilor puse la dispoziţie de grupul BAT, care e la rândul său stimulat prin intermediul diferitelor politici administrative şi industriale ale guvernului chinez. Serviciul de credit Zhima al Alibaba nu a declarat public că face parte a sistemului naţional de credite, însă calculează cu toate acestea bonitatea unei persoane pe baza unor date precum achiziţiile persoanei respective sau prietenii pe care îi are pe reţeaua socială Alipay. În 2015, directorul de tehnologie al Zhima Credit a declarat public că persoanele care cumpără scutece au un „comportament responsabil”, în vreme ce cei care joacă jocuri video în exces vor fi penalizaţi.116

Să ne aducem aminte că în introducere am vorbit despre baza de date cloud a poliţiei din China, construită pentru a monitoriza şi urmări persoanele cu tulburări psihice, cele care au criticat guvernul în public şi cele care fac parte din minorităţi etnice. Programul integrat de operaţiuni foloseşte inteligenţa artificială pentru a detecta abateri de la tiparele normale de comportament, cum ar fi traversarea prin locuri interzise. Sistemul de credite sociale din China notează şi ierarhizează cetăţenii pe baza comportamentului acestora; sistemele de luare a deciziilor bazate pe inteligenţă artificială folosesc aceste scoruri

pentru a stabili cine se poate împrumuta de la bancă, cine poate călători şi chiar şcoala la care se pot înscrie copiii acestora.

Robin Li, unul dintre fondatorii Baidu, susţine că pentru poporul chinez, intimitatea nu este o valoare primordială, aşa cum este pentru occidentali. „Poporul chinez este mai deschis sau mai puţin sensibil când vine vorba de viaţa privată”, le-a spus Li celor aflaţi în public la Forumul de Dezvoltare de la Beijing. „Dacă au de ales între intimitate şi confort, siguranţă şi eficienţă, ei vor prefera în multe situaţii să facă acest schimb”. 117 Sau poate că e vorba mai mult de repercusiuni.

Sunt de părere că sistemul chinez de credite sociale nu urmăreşte neapărat consolidarea puterii partidului comunist, sau că nu e o modalitate complicată de a dobândi un avantaj strategic faţă de cei care dezvoltă inteligenţa artificială în Occident. Acest sistem presupune mai degrabă exercitarea unui control total asupra economiei globale. La începutul anului 2018, preşedintele Xi a declarat pentru agenţia naţională de presă Xinhua că „strângând cureaua şi din dinţi, vom construi, două bombe şi un satelit”, o referire la un program de înzestrare militară din epoca maoistă. „Asta pentru că am ajuns la performanţa maximă a sistemului socialist. Ne-am unit eforturile pentru a realiza lucruri măreţe. Următorul pas este să obţinem aceleaşi rezultate prin ştiinţă şi tehnologie. Trebuie să renunţăm la iluzii şi să ne bazăm doar pe noi înşine. 118

Xi respinge noţiunea de economie de piaţă, internet liber şi ecosistem divers în care pot coexista idei concurente şi complementare. Economia naţională a Chinei, puternic centralizată, se izolează de concurenţă. Ea funcţionează pe premisa unor reţele alternative, în care regulile internetului se aplică în funcţie de locul unde este situat utilizatorul. Prin intermediul politicilor de siguranţă a datelor, exprimarea liberă este îngrădită, impunându-se o cenzură centralizată în toate domeniile celei de-a treia ere

a calculatoarelor: infrastructura internetului, traficul global de date, şi aparatura se supun din ce în ce mai mult restricţiilor impuse de Beijing. La o conferinţă din 2016, Xi a declarat că guvernul intenţionează să decidă discreţionar cum să protejeze reţelele, aparatura şi datele.119

Guvernul va exercita acest control puternic oferindu-le partenerilor din Noul Drum al Mătăsii infrastructura şi proiecte tehnologice pilot. Tanzania a fost aleasă să participe în etapa pilot de implementare – şi, poate că nu e o coincidenţă, această ţară a adoptat acum multe dintre politicile chineze referitoare la date şi informatică. Guvernul Tanzaniei beneficiază de asistenţă tehnică din partea omologilor chinezi, iar un înalt oficial din Tanzania a afirmat că „prietenii noştri chinezi au reuşit să blocheze aceste site-uri în propria ţară şi să le înlocuiască cu platforme echivalente care sunt sigure, populare şi constructive”. 120 Acelaşi lucru se petrece şi în alte ţări din Africa. Şi Vietnamul a adoptat legi draconice în privinţa securităţii cibernetice, după model chinezesc. În luna iunie 2018, India avea în vedere să adopte reglementări similare celor chineze referitoare la stocarea locală a datelor şi obţinerea unor tehnologii naţionale de securizare a datelor.121

Cum ar fi dacă China ar începe să-şi influenţeze partenerii din Noul Drum al Mătăsii, iar unul dintre principalele exporturi ar deveni sistemul naţional de credite sociale? E uşor de imaginat cum regimurile autocrate, cum ar fi Turcia sau Rwanda, ar putea cumpăra această tehnologie de la guvernul chinez. Cum rămâne însă cu alte ţări, ca Brazilia sau Austria, care au cedat tentaţiei populiste şi în care în prezent sunt la putere lideri naţionalişti? Cum ar fi dacă o agenţie guvernamentală din *ţara voastră* ar fi inspirată sau forţată să adopte un sistem de credite sociale, care să vă monitorizeze fără acordul vostru explicit? Aţi şti că aveţi un scor, sau că vă aflaţi pe o listă de persoane suspecte?

Cum ar fi dacă companiile străine ar fi împărţite în

categorii şi ar primi fie un tratament preferenţial sau Îi s-ar interzice să facă afaceri în China – sau chiar între ele? Pe măsură ce economia chineză creşte, ce s-ar întâmpla dacă această putere şi influenţă s-ar răspândi prin internet, prin aparatele pe care le folosim şi chiar prin inteligenţa artificială însăşi?

Cum ar fi dacă China ar construi un sistem de credite sociale pentru persoanele din afara ţării, folosind date pe care le extrage pe gratis din internetul liber şi reţelele de socializare occidentale? Cum ar fi dacă s-ar folosi de toate datele pe care le lăsaţi în urma voastră după ce vizitaţi Marele Zid sau Oraşul Interzis? Cum rămâne cu toate atacurile cibernetice de care auzim periodic, în care au loc breşe de securitate care-şi au originea în reţele din China?

Există şi un alt motiv pentru care ar trebui să ne îngrijoreze planurile Chinei, şi aici revenim la incubatorul triburilor inteligenţei artificiale: sistemul educaţional. China atrage constant profesori şi cercetători din centre din Canada şi Statele Unite, oferindu-le pachete de relocare atractive. Există deja o penurie de specialişti experimentaţi în domeniul învăţării automate şi stocării datelor. Acest exod al creierelor va crea în curând o lipsă de talente în Occident. Aceasta este de departe cea mai inteligentă mutare a Chinei, deoarece va priva Occidentul de capacitatea de a rămâne competitiv în viitor.

Reţeaua de recrutare chineza atrage cercetătorii înapoi în ţară ca parte a Planului celor O Mie de Talente. Expansiunea rapidă a grupului BAT a generat o cerere de oameni calificaţimajoritatea lor fiind absolvenţi în Statele Unite şi care lucrează pentru universităţile şi companiile din America. Acest plan guvernamental urmăreşte atragerea specialiştilor şi profesorilor, oferindu-le stimulente financiare considerabile (atât de ordin personal cât şi pentru proiecte de cercetare) dar şi ocazia de a se integra într-un mediu lipsit de constrângerile administrative şi juridice întâlnite în Statele Unite. Peste 7.000 de oameni au fost atraşi până acum în acest

program, primind un bonus la semnare din partea guvernului în sumă de 1 milion de yuan (aproximativ 151.000 de dolari), un buget personal de cercetare de 3 – 5 milioane de yuan (467.000 – 778.000 de dolari), subvenţii pentru locuinţă şi educaţie, norme de hrană, pachete de relocare, asistenţă pentru ca rudele să-şi găsească un loc de muncă şi chiar bilete de avion pentru ca familia să poată să se reunească.122 Toţi cei care revin în ţară îşi pun – direct sau indirecttalentele în slujba grupului BAT.

Triburile Americii: Mafia G

Dacă inteligenţa artificială este cursa spaţială a Chinei, China are un avans considerabil. În ultimii ani, în timp ce în domeniu s-au făcut descoperiri importante, administraţia Trump a redus sumele alocate cercetării, a împrăştiat zvonuri neadevărate despre impactul inteligenţei artificiale asupra forţei de muncă, a îndepărtat aliaţii strategici ai Americii şi a sâcâit în mod repetat China prin impunerea de tarife vamale.

Ne vom da în curând seama că politicienii noştri nu au o strategie coerentă pentru inteligenţa artificială, sau pentru viitorul nostru pe termen lung. În locul acestei strategii stau oportunismul şi goana după succesul de piaţă. Membrii celor Nouă Titani din Statele Unite ar putea fi companii de succes în mod individual, însă ei nu fac parte a unui demers concertat de a consolida şi centraliza puterea economică şi militară în Statele Unite. Ei nici nu ar putea – sau nu ar trebui – să fie de acord cu un astfel de plan.

Istoria membrilor americani e deja bine cunoscută, însă schimbările semnificative care vor avea loc în relaţia dintre aceste companii, datele voastre şi aparatele pe care le folosiţi sunt mai puţin cunoscute.

Membrii americani-Google, Microsoft, Amazon, Facebook, IBM şi Apple sunt companii inovatoare, inventive care au pus în mare măsură bazele progreselor în inteligenţa artificială. Ei funcţionează ca o *mafie* în sensul cel mai pur (dar nu peiorativ) al cuvântului:

constituie o reţea foarte strânsă de oameni cu interese şi provenienţe similare care colaborează în cadrul unui domeniu care ne va afecta viitorul. În momentul de faţă, Google are cea mai mare influenţă în domeniul inteligenţei artificiale, al guvernului şi al vieţii noastre de zi cu zi, aşa că le vom spune companiilor americane Mafia G. Nu e de mirare că China s-a grăbit să le emuleze succesul, sau că li s-a interzis în bună măsură să opereze în ţara respectivă. Aceste companii nu au pornit de la inteligenţa artificială ca element de bază al afacerilor, dar toate aceste şase companii s-au concentrat în ultimii trei ani în mod special asupra aplicaţiilor comerciale ale inteligenţei artificiale, prin intermediul cercetării, al parteneriatelor şi al unor noi produse şi servicii.

În China, guvernul este cel care controlează BAT. în Statele Unite, Mafia G are o putere şi o influenţă semnificativă asupra guvernului, şi pentru că în America există o economie de piaţă, dar şi datorită aversiunii culturale puternice împotriva intervenţiilor guvernamentale în mediul de afaceri. Există şi un alt motiv pentru care Mafia G are o influenţă atât de mare-membrii ei au fost ignoraţi de către politicienii de la Washington. În timp ce Xi îşi consolida poziţia şi lansa planurile pentru orizontul 2030, asistentul adjunct pentru politici tehnologice al preşedintelui Trump, Michael Kratsios, declara în faţa unui grup de lideri din domeniul tehnologiei reuniţi la Casa Albă că pentru Statele Unite este de preferat ca firmele din Silicon Valley să-şi decidă singure prioritătile, fără intervenţia statului.123

Există un dezechilibru al puterii Unite pentru că guvernul Statelor nu a fost capabil să creeze reţelele, bazele de date şi infrastructura de care are nevoie pentru a funcţiona, astfel că are nevoie de Mafia G. Spre exemplu, există şanse ca Amazon să ofere servicii de cloud computing către guvernul american în valoare de 4, 6 miliarde de dolari în 2019 – asta în timp ce compania privată spaţială a lui Jeff Bezos, Blue Origin, îşi va oferi

probabil sprijinul către NAŞA şi Pentagon în cadrul unor viitoare misiuni spaţiale. În Statele Unite, guvernul depinde de Mafia G, şi din cauză că suntem o economie de piaţă care are legi şi reglementări menite să protejeze mediul de afaceri, companiile din Silicon Valley exercită o influenţă considerabilă. Să mă fac înţeleasă: nu neg rolul Mafiei G sau că membrii ei ar fi companii profitabil şi de succes. Nici nu cred că faptul că câştigă mulţi bani le dă o aură negativă. Mafia G nu ar trebui constrânsă sau reglementată când vine vorba de obţinerea profitului, atât timă cât nu încalcă alte legi.

Această oportunitate are însă şi costuri. Există o presiune uriaşă ca Mafia G să construiască aplicaţii practice, vandabile ale inteligenţei artificiale, şi cât mai repede cu putinţă. În lumea digitală, investitorii s-au obişnuit cu profituri consistente şi rapide. Dropbox, o platformă de stocare digitală a ajuns la o valoare de piaţă de 10 miliarde de dolari după doar 6 ani de la lansare. Firma de investiţii din Silicon Valley Sequoia Capital a investit cu o participaţie de 20% la subscripţia iniţială a acţiunilor Dropbox, o participaţie care acum valorează 1, 7 miliarde de dolari.124 în Silicon Valley, companiile noi care ajung la o valoare de peste 1 miliard de dolari sunt cunoscute sub numele de „unicorni”, iar cu o valoare de 10 ori mai mare, Dropbox este un „decacorn”. În 2018 erau suficienţi unicorni şi decacorni cât pentru o grădină zoologică în Silicon Valley, iar câţiva dintre ei au parteneriate cu Mafia G, printre care aş aminti SpaceX, Coinbase, Peloton, Credit Karma, Airbnb, Palantir şi Über. Un aflux rapid de bani atrage după sine şi aşteptări mărite ca produsul sau serviciul să recupereze investiţia iniţială, fie printr-o adoptare generalizată, o achiziţie efectuată de o firmă mai mare sau un renume de piaţă.

Aveţi o relaţie specială cu Mafia G, chiar dacă nu folosiţi produsele lor bine cunoscute. Teoria celor „şase etape de distanţă” reprezintă o modalitate matematică de a explica felul în care suntem cu toţii înrudiţi. Între voi şi

toţi cunoscuţii voştri există o distanţă de un pas, o distanţă de doi paşi între voi şi cunoscuţii cunoscuţilor şi tot aşa. Între voi şi Mafia G sunt surprinzător de puţini paşi, chiar şi dacă nu folosiţi internetul.

Două treimi dintre adulţii din Statele Unite folosesc Facebook, 125 iar majoritatea utilizatorilor accesează reţeaua de socializare cel puţin o dată pe zi, ceea ce înseamnă că şi dacă nu o folosiţi, cineva din anturajul vostru o va face. Există cel mult unul sau doi paşi între voi şi Facebook, chiar şi dacă nu aţi „apreciat” niciodată o postare, şi chiar dacă v-aţi şters contul. Aproape jumătate din gospodăriile americane folosesc Amazon Prime, aşa că între voi şi Amazon există unul până la trei paşi distanţă.126 Dacă aţi fost la doctor în ultimii zece ani, între voi şi Microsoft şi IBM e doar un pas distanţă. 95% dintre americani au telefoane inteligente, 127 ceea ce înseamnă că între voi şi Apple sau Google e doar un pas.

Prin simplul fapt de a fi trăit în ultimele două decenii, aţi generat date pentru Mafia G, chiar şi dacă nu le folosiţi produsele şi serviciile. Motivul este că avem un număr uriaş de aparate inteligente care generează date: telefoanele noastre mobile, aparatele GPS, sistemele de comandă vocală inteligente, televizoarele şi aparatele de înregistrat filme, camerele de supraveghere, brăţările de fitness, senzorii fără fir din grădină şi echipamente de fitness interconectate – şi asta pentru că o proporţie covârşitoare din comunicaţiile, cumpărăturile, munca şi viaţa zilnică se petrece pe platformele Mafiei G.

În Statele Unite, firmele pot accesa contra cost aceste date în scopuri comerciale sau pot optimiza funcţionalităţile diferitelor sisteme pe care le folosim. Putem cumpăra pe multe site-uri folosind adresa de e-mail şi cardul de credit pe care le-am pus pe tavă celor de la Amazon. Ne putem autentifica pe multe portaluri folosind contul de Facebook. Capacitatea de a folosi datele pe care le deţine Mafia G pentru alte servicii este legată de datele pe care le generămsub forma pozelor, înregistrărilor audio

şi video, informaţiilor biometrice, ale traficului de internet şi aşa mai departe. Toate datele noastre sunt stocate în „cloud”, un termen în vogă care descrie programele şi serviciile care sunt stocate pe internet şi nu pe propriile voastre aparate. Poate că nu e de mirare că principalii furnizori de servicii cloud sunt Google, Amazon, Microsoft şi IBM.

Accesaţi serviciile cloud direct (când creaţi documente Google partajate) sau indirect (atunci când telefonul vostru creează automat o copie de siguranţă a fotografiilor pe serverele companiilor). Dacă aveţi un iPhone sau un iPad, folosiţi serviciul cloud al Apple. Dacă aţi accesat site-ul healthcare.gov în Statele Unite, aţi accesat serviciul cloud al Amazon. Dacă copilul vostru a avut o petrecere de zi de naştere Build-A-Bear la mall, ea a fost organizată prin serviciul cloud al Microsoft. În ultimul deceniu, serviciile cloud au luat amploareuna atât de mare încât nici nu ni se mai par remarcabile sau interesante sau stimulante din punct de vedere tehnic. Ele există pur şi simplu, la fel ca electricitatea sau apa curentă. Nu ne gândim mai deloc la ele decât când nu le mai putem accesa.

Cu toţii generăm date şi folosim serviciile cloud, având o încredere oarbă în triburile inteligenţei artificiale şi sistemele comerciale create de acestea. În Statele Unite, datele noastre spun mult mai multe despre noi decât codul de asigurări sociale pe care am fost învăţaţi să-l păzim cu sfinţenie. Pe baza numărului de asigurări sociale, o persoană poate să deschidă un cont bancar sau să solicite un credit pentru cumpărarea unei maşini. Cu datele stocate în cloud, Mafia G ar putea teoretic să ştie dacă ascundeţi o sarcină, dacă angajaţii cred că sunteţi un patron incompetent sau dacă aveţi o boală terminală-iar inteligenţa artificială a Mafiei G ştie probabil toate aceste lucruri cu mult înainte ca voi să le aflaţi. Perspectiva divină pe care o are Mafia G asupra vieţilor noastre nu e neapărat un lucru rău. Există multe feluri prin care minarea datelor

noastre personale pentru a afla informaţii noi ar putea avea ca rezultat o viaţă mai fericită şi mai sănătoasă pentru toţi.

Oricât de puternice ar fi serviciile cloud şi inteligenţa artificială a Mafiei G, ele au încă anumite limitări: aparatura. Arhitectura inteligenţei artificiale din zilele noastre a fost suficient de bună pentru a construi produse cu inteligenţă artificială limitată, cum ar fi filtrul de mesaje nedorite al Gmail, sau serviciul Apple de convertire a mesageriei vocale în text. Această arhitectură va trebui însă adaptată la inteligenţa artificială generală, o etapă mai dificilă care se întrevede în acest moment la orizont. Pentru a putea trece la următorul nivel va fi nevoie de o aparatură special creată pentru inteligenţa artificială.

Motivul este legat de un element amintit anterior de John von Neumann, informaticianul care a formulat teoria care stă la bazele arhitecturii calculatoarelor moderne. Trebuie să vă aduceţi aminte că pe vremea lui von Neumann, calculatoarele primeau programe şi date separate pentru a fi procesate în această arhitectură programele şi datele erau stocate simultan în memoria aparatului. Această arhitectură este folosită şi în zilele noastre pentru laptopuri şi computere de birou, iar datele se mişcă între procesor şi memorie. Dacă nu aveţi suficientă putere de calcul, aparatul se va încinge sau veţi primi un mesaj de eroare sau calculatorul se va bloca. Problema este cunoscută sub numele de „gâtul de sticlă von Neumann”. Indiferent cât de repede ar putea procesorul să prelucreze datele, memoria programului şi cea a datelor provoacă acest gât de sticlă care limitează rata de transfer. Mai toate calculatoarele din ziua de azi sunt bazate pe această arhitectură von Neumann, iar problema constă în faptul că procesoarele din zilele noastre nu pot rula programele mai rapid decât pot prelua instrucţiunile şi datele din memorie.

Gâtul de sticlă constituie o problemă mare a inteligenţei artificiale. Dacă vorbiţi cu sistemele Alexa sau

Google Home, vocea voastră este înregistrată, convertită în date şi apoi transmisă serverului din cloud pentru o reacţie – dată fiind distanţa fizică între voi şi diferitele locuri unde se găsesc serverele implicate, e uluitor că Alexa poate să reacţioneze în doar o secundă sau două. Pe măsură ce tot mai multe aparate capătă o componentă legată de inteligenţa artificială – sub forma telefoanelor inteligente cu senzori biometrici, camerelor de supraveghere care ne pot recunoaşte faţa, maşinilor care se conduc singure, sau a roboţilor precişi care pot doza medicamenteo întârziere de procesare de doar o secundă sau două poate provoca un rezultat catastrofal. O maşină autonomă nu poate contacta serverul cloud pentru fiecare acţiune, pentru că sunt mult prea mulţi senzori care ar trebui să transmită datele către serverul cloud pentru a fi procesate.

Singura soluţie o reprezintă deplasarea puterii de calcul mai aproape de sursa datelor, ceea ce va conduce la reducerea timpului de răspuns şi a încărcării reţelelor. Această nouă arhitectură se numeşte „edge computing” şi reprezintă evoluţia inevitabilă a aparatelor dotate cu inteligenţă artificială şi a arhitecturii sistemelor. Pentru ca inteligenţa artificială să progreseze la următoarea etapă, aparatura trebuie să ţină pasul cu ea. În loc să ne întâlnim cu Mafia G pe serverul cloud, unde încă mai avem o capacitate de a modifica permisiuni şi setări, vom fi nevoiţi să-i invităm să controleze *toate* aparatele pe care le folosim. Aceasta înseamnă că în următorul deceniu, restul ecosistemului inteligenţei artificiale va funcţiona doar pe baza câtorva sisteme ale Mafiei G. Toate companiile nou înfiinţate şi firmele marginaleca să nu mai vorbim de oamenii ca noivor trebui să accepte o nouă ordine şi să jure credinţă unei mâini de furnizori de servicii care vor furniza sistemele de operare pentru viaţa de zi cu zi. Odată ce datele, aparatele, maşinile şi serviciile vor fi interconectate, veţi deveni captivi. Pe măsură ce cumpăraţi mai multe bunuri – telefoane mobile, frigidere cu

conexiune la Internet, căşti inteligent – veţi constata că Mafia G vă oferă sistemul de operare pentru viaţa zilnică. Omenirea primeşte o ofertă pe care nu o poate refuza.

Calculele legate de învăţarea profundă necesită aparatură specializată pentru că au nevoie de multă putere de calcul. Deoarece acestea preferă optimizarea în favoarea preciziei şi se compun în esenţă din operaţiuni complexe de algebră liniară, este de înţeles că o nouă arhitectură a reţelelor neurale ar provoca o eficienţă mărită, dar, şi mai important, o accelerare a procesului de design şi lansare pe piaţă. Cu cât echipele de cercetare pot construi şi testa modele reale, cu atât se vor apropia mai mult de aplicaţii practice ale inteligenţei artificiale. Spre exemplu, în prezent e nevoie de săptămâni sau de luni pentru a pregăti un model complex de vizualizare digitalăiar rezultatul final ar putea doar să dovedească nevoia unor ajustări suplimentare, ceea ce înseamnă că munca trebuie reluată de la capăt. O aparatură mai bună ar presupune ca aceste modele să poată fi pregătite în câteva ore, sau câteva minute, ceea ce ar conduce la noi descoperiri săptămânale, sau chiar zilnice.

Acesta este motivul pentru care Google şi-a creat propriile circuite, denumite Unităţi de Procesare prin Tensori (UPT). Aceste circuite pot funcţiona cu structura de învăţare profundă prin inteligenţă artificială, Tensor Flow. În luna iunie 2018, Tensor Flow era principala platformă de învăţare automată de pe GitHub, care este cea mai mare platformă de pe internet din lume în care programatorii îşi depozitează softul. Tensor Flow a fost descărcată de peste 10 milioane de ori de către programatori din 180 de ţări, la momentul redactării acestei cărţi existând 24.500 de copii active ale platformei.128 Google a completat platforma cu produse suplimentare, cum ar fi TF-GAN (o bibliotecă de module pentru reţele neurale opuse) e şi programul Tensor Flow de detectare a obiectelor (care îi ajută pe programatori să creeze modele mai exacte de învăţare pentru vizualizarea

digitală). Circuitele UPT sunt deja folosite în serverele Googleele stau la baza modelelor de învăţare profundă care se activează cu fiecare căutare de pe Google.

Google avea motive întemeiate să încerce să achiziţioneze GitHub, o platformă cu 28 de milioane de programatori din lumea întreagă şi cu legături puternice cu ceilalţi membri ai celor Nouă Titani. În luna iunie 2018, Google a fost depăşită de – cine credeţi?

— Microsoft.129

Facebook a colaborat cu Intel pentru dezvoltarea unui chip de inteligenţă artificială pentru propria activitate de cercetare de care compania avea nevoie pentru a accelera procesul de experimentare. Apple şi-a creat propriul circuit „cu motor neural” pe e Reţele neurale la care învăţarea se produce din interacţiunea între reţele opuse una celeilalte. (n.t.) care îl foloseşte în cadrul telefoanelor iPhone X, iar Microsoft a dezvoltat circuite de inteligenţă artificială pentru casca de realitate virtuală Hololens şi platforma de cloud computing Azure. Membrii BAT îşi construiesc de asemenea propriile circuite: Alibaba a început un proces intens de recrutare în Silicon Valley, căutând „arhitecţi pentru circuite de inteligenţă artificială” 130, iar în 2018 şi-a lansat propriile circuite – Ali-NPU – pe care oricine le poate folosi prin intermediul serviciului public cloud.

Anticipând faptul că în viitor cerinţele de performanţă vor creşte, IBM a dezvoltat acum câţiva ani circuitul neuromorfologic True North, lucrând deja la un nou tip de aparate care ar putea spori eficienţa reţelelor neurale de peste 100 de ori. Ca să înţelegeţi despre ce e vorba, e ca o diferenţă de la un abac la maşina de teleportare din Star Trek. Această nouă generaţie de circuite foloseşte două tipuri de sinapse, una pentru memoria pe termen lung, şi alta pentru calcule pe termen scurt.

Vorbim despre un echivalent al zilelor noastre al întrebării „eşti genul de om care foloseşte un PC sau un

Mac?” dus la extrem. Majoritatea acestor circuite funcţionează pe baza unor structuri pe care cei Nouă Titani le definesc a fi libere de drepturi de autoreeea ce înseamnă că dezvoltatorii pot să acceseze, să folosească şi să îmbunătăţească gratuit structurile. Circuitele fizice sunt însă protejate, iar serviciile sunt bazate pe abonamente. În practică, aceasta înseamnă că odată ce o aplicaţie este construită pentru o anumită structură, ea va fi deosebit de dificil de adaptat pentru o altă structură. Acesta este modul prin care triburile inteligenţei artificiale recrutează noi aderenţi-iar ritualul de iniţiere este sărutarea inelului unei structuri a Mafiei G.

Căutând să comercializeze inteligenţa artificială, Mafia G recrutează programatori într-o manieră creativă. În luna mai 2018, Google şi platforma de cursuri online Courseră au lansat o nouă specializare dedicată învăţării automate. Pentru a participa, trebuie însă să folosiţi Tensor Flow. Cursul cu cinci părţi, la finalul căruia obţineţi un certificat de absolvire, este descris ca o modalitate accesibilă oricui de a învăţa despre învăţarea automată şi reţelele neurale. Cursanţii au nevoie de date şi structuri din lumea reală, aşa că se bazează pe structura pusă la dispoziţie de Google.

Aparatura face parte din strategia pentru inteligenţă artificială a Mafiei G, care are legături cu guvernul, diferite de cele despre care am vorbit în cazul Chinei, dar care ar trebui să vă provoace îngrijorări similare, chiar şi dacă nu sunteţi cetăţean al Statelor Unite.

Motivul este că în Statele Unite, inteligenţa artificială este pusă în slujba a trei stăpâni: guvernul, lumea finanţelor de pe Wall Street şi companiile din Silicon Valley. Oamenii care redactează politicile guvernamentale sau reglementările sunt tehnocraţi de carieră, care au tendinţa să-şi păstreze zeci de ani slujbele. Cei care deţin însă controlul – preşedintele sau şefii marilor agenţii guvernamentale (Federal Communications Commission, Departamentul de Justiţie şi alţii) se schimbă din funcţie o

dată la câţiva ani. În domeniul inteligenţei artificiale nu există o direcţie clară la nivel naţional.

Doar recent a început să fie luată în calcul mai clar poziţia Chinei şi planurile acesteia în domeniul inteligenţei artificiale, în principal din cauza faptului că preşedintele Xi a dat publicităţii planul strategic pe termen lung asupra inteligenţei artificiale şi folosirii datelor. În Statele Unite există o autoritate denumită Cornitetul pentru Investiţii străine. Acesta este o comisie compusă din reprezentanţi ai puterii şi opoziţiei condus de secretarul Trezoreriei şi compus din membri ai Departamentelor Trezorerie, Justiţiei, Energiei, Apărării, Comerţului, de Stat şi al Siguranţei Naţionale. Misiunea lor este de a analiza şi investiga tranzacţiile comerciale care ar putea periclita siguranţa naţională. Acest comitet a blocat achiziţia companiei Qualcomm, un fabricant de circuite din San Diego, de către firma Broadcom din Singapore. Acelaşi comitet a respins şi achiziţia companiei Money Gram din Dallas de către compania de plăţi electronice Ani Financial, deţinută de Alibaba. La momentul scrierii cărţii, comitetul nu acordă o atenţie specială inteligenţei artificiale, deşi s-a propus ca puterile sale să fie sporite pentru a preîntâmpina expansiunea Chinei prin cumpărarea companiilor americane.

Între timp, în Silicon Valley a devenit o practică curentă ca angajaţii să treacă de la o compania la alta, în timp ce liderii tribului îşi împart timpul între Mafia G şi au o stabilitate mai mare a slujbei. Inteligenţa artificială continuă să progreseze în timp ce mantra tribului – *construieşte acum*, *cere-ţi scuze mai* târziudevine din ce în ce mai puternică. Ani de zile Google a scanat şi catalogat cărţi protejate de drepturi de autor fără să ceară în prealabil permisiunea autorilorcompania a fost acţionată în justiţie de edituri şi autori. Google a fotografiat casele şi cartierele noastre şi le-a publicat pe Google Maps, fără să ne ceară permisiunea. (Oamenii sunt evitaţi, pe cât posibil, iar chipurile lor sunt voalate). Apple încetineşte deliberat

performanţa telefoanelor mai vechi când lansează modele noi, apoi îşi cere scuze. După scandalul Cambridge Analytica, directorul general al Facebook Mark Zuckerberg a publicat o dezminţire pe propriul zid de Facebook, scriind „Celor pe care i-am dezamăgit, le cer iertare şi voi încerca să fiu mai bun. Cer iertare pentru felul în care munca mea a contribuit la divizarea oamenilor în loc să ne aducă împreună”.

Prin urmare, Mafia G are tendinţa de a acţiona cu rapiditate în salturi evolutive până când se petrece ceva rău, apoi se implică guvernul. Politicile Facebook de securitate a datelor nu au atras atenţia guvernului decât în momentul în care un fost angajat Cambridge Analytica a rupt tăcerea, explicând uşurinţa cu care au fost adunate şi vândute datele noastre. În 2016, ca urmare a unui atac armat petrecut la San Bernardino, California, guvernul federal a încercat să oblige Apple să ofere autorităţilor o cheie de decriptare a iPhone-ului atacatorului. Agenţiile guvernamentale şi oamenii legii susţineau că decodarea telefonului şi transmiterea datelor private era în interesul publicului, în timp ce apărătorii dreptului la intimitate susţineau că ar viola drepturile civile. În cele din urmă, autorităţile au reuşit să spargă telefonul fără ajutorul Apple, aşa că nu am apucat să vedem care tabără a avut dreptate. În Statele Unite intimitatea este la mare preţ, dar nu avem legi clare, care să lămurească situaţia datelor personale în secolul 21.

În vara anului 2018, personalul de la cabinetul senatorului Mark Warner (D-VA) a lansat în public o serie de propuneri menite să înfrâneze avântul giganţilor din domeniul tehnologiei. Printre acestea se numără crearea unei legislaţii asemănătoarea cu reglementările GDPR din Europa sau propunerea ca platformele de internet să fie considerate persoane de încredere care trebuie să urmeze un anumit cod de conduită, asemănător cu al firmelor de avocatură.131 La doar câteva luni după aceea, directorul general al Apple, Tim Cook, a postat pe Twitter câteva

gânduri despre viitorul intimităţii, coloşii din tehnologie şi America. Pe 24 octombrie el a scris ca protecţia intimităţii utilizatorilor ar trebui să reprezinte principala prioritate a companiilor. „Companiile trebuie să recunoască faptul că datele aparţin utilizatorilor şi ar trebui să înlesnim capacitatea oamenilor de a a obţine o copie a datelor lor personale, precum şi de a le corecta sau de a le şterge”, scria el, continuând, „Oricine are dreptul ca datele proprii să fie în siguranţă”. 132 Simţind că în Statele Unite reglementările în domeniul devin plauzibile, Apple şi-a promovat serviciile de protecţie a datelor precum setările de intimitate incluse în sistemele de operare pentru terminale mobile şi computere.

Suntem de acord să fim supravegheaţi constant în schimbul serviciilor de care beneficiem. Acest troc permite Mafiei G să facă profit, pentru a-şi putea îmbunătăţi şi extinde paleta de servicii pe care ni le oferă, fie în calitate de consumatori individuali, fie ca firme, universităţi, organizaţii neguvernamentale sau agenţii ale statului. Avem de-a face cu un model de afaceri bazat pe capitalismul de supraveghere. Şi, dacă e să fim sinceri, e un sistem pe care în Statele Unite l-am acceptat, altminteri am fi încetat de mult să mai folosim servicii ca Gmail, Microsoft Outlook sau Facebook. Pentru a putea funcţiona cum trebuie, ei au nevoie de datele noastre, care sunt extrase, procesate şi ambalate. Presupun că folosiţi cel puţin unul dintre produsele sau serviciile oferite de către Mafia G. Eu folosesc zeci de astfel de produse şi sunt perfect conştientă care este preţul lor real.

Ceea ce vreau să spun este că în curând nu vom încredinţa Mafiei G doar datele noastre. Pe măsură ce trecem de la o inteligenţă artificială limitată la una mai vastă, capabilă să ia decizii complexe, vom invita aceste companii direct în frigiderele, dulapurile şi maşinile noastre, şi în ochelarii, brăţările şi căştile interconectate pe care le vom purta în curând. Astfel, Mafia G va putea să automatizeze muncile repetitive, să ne ajute să luăm

decizii şi să ne ajute să gândim mai uşor. Între noi şi Mafia G nu va mai exista niciun pas de distanţă. Politicienii nu vor mai putea invoca nicio autoritate reală odată ce întreaga noastră existenţă va fi strâns legată de aceste companii. La ce am putea însă renunţa, în schimb?

Cei nouă titani – grupul BAT (Baidu, Alibaba şi Tencent) şi Mafia G (Google, Microsoft, Amazon, Facebook, IBM şi Apple) – dezvoltă instrumentele şi mediul care vor sta la baza viitorului inteligenţei artificiale. Aceste companii fac parte din tribul inteligenţei artificiale, format în universităţi în care se induc idealuri şi obiective comune, care se rigidizează şi mai mult odată cu intrarea absolvenţilor pe piaţa muncii. Domeniul inteligenţei artificiale nu este unul static. Pe măsură ce inteligenţa artificială îngustă evoluează către o inteligenţă artificială generală, cei Nouă Titani dezvoltă noi tipuri de sisteme de calcul şi recrutează programatori care devin prizonieri ai structurilor lor.

Modelul consumist al inteligenţei artificiale sub forma întâlnită în Statele Unite nu este în mod obligatoriu rău, aşa cum nici modelul chinez, bazat pe intervenţia centralizată a statului nu este. Inteligenţa artificială nu este neapărat dăunătoare societăţii. Cu toate acestea, Mafia G se compune din companii care urmăresc profiturile, companii listate la bursă care trebuie să dea socoteală acţionarilor, indiferent de intenţiile altruiste ale conducătorilor şi angajaţilor lor. În China, grupul BAT se supune guvernului chinez, care a stabilit deja prioritătile cetăţenilor. Ceea ce aş vrea să ştiu – şi ceea ce ar trebui să vă intereseze şi pe voieşte ce e mai bine pentru toată omenirea? Pe măsură ce inteligenţa artificială evoluează, cum se vor reflecta deciziile pe care le luăm astăzi în deciziile pe care maşinile le vor lua în numele nostru în viitor?

CAPITOLUL TREI

O MIE DE TĂIETURI CU HÂRTIA: CONSECINŢELE NEINTENŢIONATE ALE

INTELIGENTEI ARTIFICIALE

„Mai întâi ne formăm obiceiurile, apoi obiceiurile ne formează pe noi. „— JOHN DRYDEN

„Tu eşti creatorul meu, dar eu sunt stăpânul tău”.

— MONSTRUL LUI FRANKENSTEIN (DE MARY SHELLEY)

Contrar tuturor poveştilor apocaliptice pe care le-aţi citit şi văzut în care inteligenţa artificială se răscoală dintr-odată şi se decide să elimine omenirea, nu va exista un singur eveniment în care tehnologia să dea greş. Ceea ce urmează să trăim cu toţii e mai degrabă ca o serie treptată de tăieturi cu hârtia. Dacă vă tăiaţi o dată la deget, e enervant, dar vă puteţi vedea în continuare de treabă. Dacă tot corpul e plin de mii de mici tăieturi, nu veţi muri, dar viaţa va deveni un chin. Micile amănunte ale vieţii-încălţatul, a mânca un taco, a dansa la nunta unui văr-nu ar mai fi alegeri. Veţi deveni nevoiţi să învăţaţi să trăiţi un alt fel de viaţă. O viaţă plină de restricţii. O viaţă cu consecinţe dureroase.

Ştim deja că etica şi diversitatea nu sunt obligatorii în universităţile unde se formează triburile inteligenţei artificiale şi nici în cadrul celor Nouă Titani, unde lucrează mai apoi triburile inteligenţei artificiale. Ştim că Mafia G funcţionează prin accelerarea acelor proiecte şi studii care au şanse de a fi vândute pe piaţă, iar grupul BAT urmăreşte îndeplinirea unui plan centralizat al guvernului chinez. Devine evident că poate nimeni – nu o agenţie mondială de reglementare (asemănătoare Agenţiei Internaţionale pentru Energia Atomică), sau un grup de universităţi sau măcar un grup de cercetătorinu se gândeşte la întrebările dificile legate de ecartul care se creează, la conflictul dintre valorile oamenilor şi valoarea economică a planurilor chineze de supremaţie în domeniul inteligenţei artificiale, sau obiectivele mercantile ale companiilor din Silicon Valley. Echilibrul dintre valorile umane şi obiectivele economice nu a constituit o prioritate

în trecut deoarece cei Nouă Titani au adus prosperitate, ne oferă servicii şi produse bune pe care le folosim cu toţii şi ne fac să ne simţim stăpâni pe domeniul nostru digital. Nu am cerut răspunsuri legate de valori pentru că momentan vieţile noastre ni se par mai bune datorită celor Nouă Titani.

Am fost însă deja tăiaţi cu hârtia de convingerile şi motivaţiile creatorilor inteligenţei artificiale. Cei Nou Titani nu se ocupă doar cu aparatura şi programele. Ei construiesc maşini gânditoare care reflectă valorile omenirii. Ecartul prezent dintre triburile inteligenţei artificiale şi oamenii obişnuiţi produce deja rezultate îngrijorătoare.

Algoritmul valorilor

V-aţi întrebat vreodată de ce sistemul inteligenţei artificiale nu e mai transparent? V-aţi întrebat ce fel de date sunt folositeinclusiv datele voastre personale pentru a ajuta inteligenţa artificială să înveţe? în ce situaţii este învăţată inteligenţa artificială să facă excepţii? Cum reuşesc să împace creatorii mercantilizarea inteligenţei artificiale cu nevoile omeneşti primare, cum ar fi intimitatea, siguranţa, sentimentul de apartenenţă, respectul de sine şi auto-actualizarea? Care sunt imperativele morale ale tribului inteligenţei artificiale? Ce cred ei că e bine şi rău? Reuşesc ei să înveţe inteligenţa artificială ce este empatia? (La drept vorbind, e util sau merită să încercăm să învăţăm inteligenţa artificială ce este empatia omenească?)

Fiecare dintre cei Nouă Titani are un set formal de valori, însă aceste coduri nu răspund la aceste întrebări. În schimb, aceste valori declarate sunt convingeri profunde care unesc, inspiră şi motivează angajaţii şi acţionarii. Valorile unei companii acţionează ca un algoritm – un set de reguli şi instrucţiuni care influenţează cultura organizaţională, stilul de conducere şi joacă un rol important în toate deciziile luate, de la şedinţele conducerii până la liniile individuale ale unui program.

Absenţa unor anumite valori afirmate este remarcabilă şi pentru că dacă ele exprimate explicit, devin greu de sesizat şi uşor de uitat.

La început, Google funcţiona pe baza unei valori simple, fundamentale: „nu fii rău”. 133 în scrisoarea de intenţie de la subscripţia publică de acţiuni în 2004, fondatorii Sergey Brin şi Larry page scriau: Eric [Schmidt], Sergey şi cu mine intenţionăm ca Google să fie un alt fel de companie, care va pune în practică valorile pe care le-a acumulat când era o companie deţinută privată şi în viitorul ei de companie listată la bursă… Ne va interesa un rezultat pe termen lung, fără să urmărim în mod special veniturile trimestriale. Vom sprijini anumite proiecte riscante dar cu potenţial mare şi vom gestiona portofoliul de proiecte… Ne vom ridica la înălţimea principiului, nu fii rău’ menţinând încrederea utilizatorilor”. 134

„Principiile de conducere” ale Amazon se regăsesc în structura de conducere, iar aceste valori se sprijină pe încredere, metrici, viteză şi rezultate. Printre principiile publicate se numără:

„Liderii pornesc de la utilizator şi se mulează pe el. Ei vor lucra intens pentru a dobândi şi a păstra încrederea clientului…” „Liderii au constant standarde înalte” pe care cei din exterior le-ar putea considera „nerezonabil de înalte”.

„Multe decizii şi acţiuni sunt reversibile şi nu necesită o analiză amănunţită. Preţuim asumarea calculată a riscurilor”.

„Fă mai mult cu mai puţin. Nu veţi câştiga puncte în plus dacă sporiţi artificial numărul de angajaţi, dimensiunea bugetului sau cheltuielile fixe”. 135

Facebook enumeră „5 valori fundamentale”, respectiv „a fi îndrăzneţ”, „a te concentra asupra impactului”, „a acţiona rapid”, „a fi transparent” faţă de activitatea companiei şi „a aduce valoare” pentru utilizatori.136 Pe de altă parte, „filosofia de management” a Tencent pune accent pe „instruirea şi încurajarea

angajaţilor pentru a obţine reuşite” pe baza „unei atitudini de încredere şi respect” şi luarea deciziilor pe baza unei formule intitulate „Integritate + Proactivitate + Colaborare + Inovare”. 137 Alibaba pune mai presus de orice „o atenţie constantă acordată nevoilor clienţilor noştri”, precum şi munca în echipă şi integritatea.138

Dacă aş desena o diagramă Venn a tuturor valorilor şi principiilor de funcţionare ale celor Nouă Titani, am observa câteva zone în care ele se suprapun. Toate aceste companii pretind angajaţilor şi echipelor să se perfecţioneze constant, să construiască produse şi servicii de care clienţii să nu se poată lipsi şi să îndeplinească aşteptările acţionarilor. Mai presus de orice, ele apreciază încrederea. Aceste valori nu sunt deosebite-ele par a fi valorile majorităţii companiilor americane.

Pentru că inteligenţa artificială are potenţialul de a avea un impact major asupra omenirii, valorile celor Nouă Titani ar trebui precizate explicitiar noi ar trebui să le pretindem un standard mai înalt decât celorlalte companii.

Ceea ce lipseşte este o formulare neechivocă prin care să se afirme faptul că omenirea reprezintă punctul central al evoluţiei inteligenţei artificiale şi că toate demersurile din acest domeniu ar trebui să se axeze pe îmbunătăţirea condiţie umane. Acest lucru ar trebui afirmat explicitiar această formulare ar trebui să se regăsească şi în alte documente ale companiilor, în şedinţele de conducere, în cadrul echipelor care lucrează în inteligenţa artificială şi în cadrul discuţiilor de marketing şi vânzări. Spre exemplu, ar trebui ca dincolo de inovaţie şi eficienţă să se discute şi despre accesibilitate. Milioane de oameni au handicapuri: au probleme de vorbire, auz, văz, tactile sau de gândire. Sau să se discute despre valorile economice, despre puterea platformelor de a creşte şi de a crea prosperitate fără să excludă anumite persoane sau grupuri.

Sau despre valorile sociale, cum sunt integritatea, diversitatea, toleranţa şi curiozitatea.

În timp ce scriam această carte, directorul general al Google, Sundar Pichai, a anunţat că Google a redactat un nou set de principii fundamentale referitoare la demersurile companiei în domeniul inteligenţei artificiale. Totuşi, aceste principii nu merg atât de departe încât să definească omenirea ca fiind punctul de referinţă pentru demersurile viitoare ale Google în acest domeniu. Comunicarea acestor principii nu constituie o realiniere strategică a valorilor fundamentale ale companiei, ci o reacţie datorată scandalului Project Maven-dar şi unui incident petrecut la începutul anului. Un grup de ingineri cu experienţă a descoperit că un proiect la care lucrau – o metodă de siguranţă pentru serviciile de tip cloud – era menită să ajute compania să câştige contracte cu armata. Amazon şi Microsoft aveau deja un grad de încredere ridicată pentru un serviciu de tip cloud dedicat guvernului, deţinând astfel autorizaţia de stocare a unor date confidenţiale. Google voia să câştige contracte bănoase cu Departamentul Apărării, iar în momentul în care inginerii au aflat, s-au răzvrătit. Ca urmare a aceste răzvrătiri, 5% dintre angajaţii Google au condamnat public proiectul Maven.139

Acesta a fost începutul unei serii de proteste care a început în 2018, când o parte a tribului şi-a dat seama că munca lor era pusă în folosul unei cauze pe care nu o sprijineau, aşa că a solicitat o schimbare. Ei au presupus că valorile lor personale erau şi cele ale companiei-şi atunci când şi-au dat seama că lucrurile nu stau aşa, au protestat. Acest incident demonstrează problemele spinoase care pot apărea atunci când membrii Mafiei G nu respectă standarde mai înalte decât companiile care oferă produse mai puţin spectaculoase.

Nu e deci de mirare că o bună parte din principiile Google legate de inteligenţa artificială vizează armele şi contractele cu armata: Google nu va crea tehnologii care să poată fi folosite militar şi al căror principal obiectiv este rănirea altor oameni, Google nu va crea sisteme de

inteligenţă artificiale care să contravină principiilor general acceptate de drept internaţional, şi aşa mai departe. „Vrem să fie clar că deşi nu elaborăm sisteme de inteligenţă artificială care să fie folosite pentru a crea arme, vom continua să colaborăm cu guvernul şi armata”, susţinea documentul.140

Spre meritul său, Google susţine că principiile sunt menite să reprezinte standarde concrete, nu doar noţiuni teoretice – iar documentul adresează problema prejudecăţilor din seturile de date. În document nu se aminteşte însă nicăieri problema transparenţei decizionale a inteligenţei artificiale, sau a căror seturi de date sunt folosite. Nu se vorbeşte de faptul că în domeniul inteligenţei artificiale lucrează pentru Google triburi omogene. Niciunul dintre aceste standarde concrete nu pune interesele omenirii mai presus de cele ale bursei.

Problema este una de transparenţă. Dacă guvernul Statelor Unite nu este capabil să construiască sistemele de care avem nevoie pentru a ne păstra siguranţa naţională, e de aşteptat ca el să apeleze la o companie care să poată face acest lucru – şi lucrurile au funcţionat în acest mod încă de la Primul Război Mondial. Am uitat prea uşor că pacea se câştigă constant şi că doar o armată bine pregătită ne garantează siguranţa şi securitatea naţională. Ministerul Apărării nu este o instituţie tiranică, care să-şi dorească super-arme bazate pe inteligenţa artificială pentru a rade de pe faţa pământului aşezări din afara ţării. Armata Statelor Unite are misiuni dincolo de uciderea oamenilor şi aruncarea în aer a obiectivelor inamice. Dacă cei care lucrează în cadrul Mafiei G nu au înţeles prea bine acest aspect, asta se datorează şi faptului că prea puţini oameni au reuşit să facă o punte între guvern şi Silicon Valley.

Ar trebui să reflectăm cu toţii la faptul că cei Nouă Titani construiesc sisteme care se bazează pe oameni, dar valorile care exprimă dorinţa noastră de a îmbunătăţi calitatea vieţii oamenilor nu sunt precizate în mod explicit.

Dacă valorile tehnologice, economice şi sociale nu fac parte din carta de valori a unei companii, este puţin probabil ca acea companie să urmărească binele omenirii în procesul de cercetare, dezvoltare şi lansare pe piaţă. Acest ecart al valorilor nu este întotdeauna evident într-o organizaţie, iar acest fapt implică un risc semnificativ atât pentru Mafia G cât şi pentru BAT, deoarece el îndepărtează angajaţii de potenţialele rezultate negative ale activităţii lor. Când indivizii şi echipele nu sunt conştienţi de acest ecart al valorilor, ei nu vor lua măsuri în privinţa unor probleme esenţiale când produsele sunt construite, testate, promovate, lansate şi vândute. Nu înseamnă că persoanele care lucrează în domeniul inteligenţei artificiale ar fi indiferentedar înseamnă că ele nu ţin cont de valorile de bază ale omenirii.

Şi aşa ajungem să ne tăiem cu hârtia.

Legea lui Conway

Informatica, la fel ca orice alt domeniu al tehnologiei, sau orice altă activitate, reflectă perspectiva asupra lumii şi experienţele echipei care se ocupă de inovaţie. Vedem aceste lucruri şi în afara tehnologiei. Permiteţi-mi să mă îndepărtez de inteligenţa artificială preţ de o clipă şi să vă prezint două exemple aparent distinete în care un mic trib de persoane poate ajunge să dobândească o putere extraordinară asupra unei populaţii.

Dacă aveţi părul drept-gros, aspru, fin, lung, scurt, subţire (sau chiar rar) – veţi avea o cu totul altă experienţă la coafor faţă de mine. Fie că mergeţi la o frizerie sau la un salon cu pretenţii, mai întâi vi se spală părul într-o mică chiuvetă, în care cineva trece cu degetele prin părul vostru. Frizerul sau stilistul foloseşte apoi un pieptene des ca să vă întindă părul în linii drepte. Dacă aveţi mult păr, stilistul ar putea folosi o perie şi un uscător de păr, trăgând din nou de fiecare şuviţă până când are forma dorită – plină de volum şi elastic, sau plată şi netedă. Dacă aveţi părul mai scurt, peria e mai mică şi timpul de uscare mai scurt, dar procesul e în linii mari acelaşi.

Părul meu e foarte creţ şi foarte des. Se încurcă uşor şi reacţionează imprevizibil la factorii de mediu. În funcţie de umiditate, de nivelul de hidratare şi de produsele cosmetice pe care le-am folosit, părul meu stă bine sau poate fi un dezastru încâlcit. La un salon obişnuit, chiar şi acelea unde nu am avut probleme, chiuveta îmi dă de furcă. Persoana care-mi spală părul are de regulă nevoie de mult mai mult spaţiu decât permite chiuveta, şi uneori părul se încâlceşte în jurul furtunului, moment în care devine dureros de descâlcit. Singura modalitate prin care un pieptene obişnuit poate trece prin părul meu e dacă părul e ud şi acoperit de un lichid alunecos, cum ar fi un balsam mai gros. (De perie nici nu poate fi vorba). Puterea unui uscător obişnuit de păr mi-ar face părul să se înnoade. Unele saloane au un capăt special care împrăştie aerul – arată ca un bol de plastic cu nişte mici protuberanţe ca nişte ardei iuţi dar pentru a-l putea folosi eficient trebuie să-mi las capul în jos, iar stilistul trebuie să se aplece ca să poată poziţiona corect uscătorul.

Aproximativ 15% dintre albi au părul creţ. Dacă adunăm şi populaţia negrilor, rezultă cam 79 de milioane de oameni, aproximativ un sfert din populaţia Statelor Unite, oameni care au probleme la frizer pentru că, putem presupune, instrumentele şi mediul au fost create de oameni cu părul drept, care nu au ţinut cont de valori sociale, cum ar fi empatia şi diversitatea, în cadrul companiilor lor.141

Acesta e un exemplu destul de inofensiv. Gândiţi-vă acum la o situaţie în care mizele sunt mult mai mari decât o simplă coafură, în luna aprilie 2017, agenţii de îmbarcare de la un zbor supraaglomerat al United Airlines de la aeroportul internaţional Chicago O'Hare au făcut un anunţ la difuzor, cerând pasagerilor să cedeze locurile lor angajaţilor companiei în schimbul unei compensaţii de 400 de dolari şi o cameră gratuită la un hotel din zonă. Nimeni nu a acceptat oferta. Angajaţii au crescut apoi oferta la 800 de dolari plus camera de hotel, dar, din nou, nimeni nu

a acceptat. Între timp, pasagerii cu prioritate la îmbarcare începuseră să se îmbarce în avion, inclusiv cei de la clasa I.

Un algoritm şi un sistem automat a ales să elimine patru oameni, printre care se număra doctorul David Dao şi soţia lui, de asemenea medic. El le-a spus personalului liniei aeriene că are pacienţi pe care trebuie să-i consulte a doua zi. În timp ce ceilalţi pasageri au acceptat să fie îndepărtaţi, Dao a refuzat să se mişte. Autorităţile l-au ameninţat pe Dao cu închisoarea dacă refuză să plece. Aţi văzut cu siguranţă ce a urmat pentru că videoclipul incidentului a devenit viral pe Facebook, YouTube şi Twitter şi a fost retransmis apoi zile la rând pe canalele de ştiri din toată lumea. Poliţiştii l-au apucat pe Dao de braţe şi l-au scos din scaun cu forţa, lovindu-l de o cotieră, distrugându-i ochelarii şi provocându-i o tăietură în zona gurii. Cu faţa plină de sânge, Dao s-a oprit dintr-odată din urlete în timp ce poliţia îl ţâra pe coridorul avionului United. Incidentul a fost traumatizant atât pentru Dao cât şi pentru ceilalţi pasageri, generând un val de reacţii negative împotriva United, care s-a soldat cu o anchetă a Congresului. Toţi au vrut să ştie: cum se poate petrece aşa ceva în Statele Unite?

Pentru majoritatea liniilor aeriene din lume, inclusiv pentru United, procedura de îmbarcare este automatizată. Southwest Airlines atribuie pasagerilor un grup (A, B sau C) şi un număr, apoi îmbarcarea se produce în ordine, sortarea fiind efectuată printr-un algoritm. Se acordă prioritate în funcţie de preţul plătit pentru bilet, statutul de călător frecvent şi momentul când a fost achiziţionat biletul. Alte companii, care folosesc locuri prealocate, efectuează îmbarcarea în grupuri bazate pe priorităţi, care sunt de asemenea atribuite în baza unui algoritm. Când vine momentul să vă urcaţi în avion, agenţii de îmbarcare urmează un set de instrucţiuni de pe un ecran – procesul este creat pentru a fi respectat cu stricteţe şi fără abateri.

Am participat la o conferinţă din industria turismului

în Houston la câteva săptămâni după incidentul celor de la United şi i-am întrebat pe directorii din industria informatică ce rol ar fi putut juca inteligenţa artificială în acest caz. Ipoteza mea era următoarea: procesul algoritmic de decizie a urmat nişte paşi prestabiliţi pentru a oferi soluţia fără să ţină cont de vreun context. Sistemul a stabilit că nu erau locuri suficiente, a calculat suma compensaţiei, iar în momentul în care situaţia nu s-a rezolvat, sistemul a ajustat compensaţia. Când pasagerul nu s-a conformat, sistemul a recomandat intervenţia forţelor de ordine. Personalul nu a făcut decât să respecte orbeşte instrucţiunile de pe ecran, ascultând automat un sistem de inteligenţă artificială care nu ţinea cont de flexibilitate, împrejurări sau empatie. Directorii din informatică, angajaţi ai altor companii decât United, nu au negat problema reală: în ziua în care Dao a fost târât afară din avion, oamenii s-au supus unui sistem de inteligenţă artificială creat de un număr mic de oameni care probabil nu au luat în considerare suficient scenariile în care sistemul ar putea fi folosit în viitor.

Instrumentele şi mediul saloanelor de coafură sau al platformelor de programe din industria aeronautică sunt exemple ale legii lui Conway, care afirmă că în absenţa unor reguli şi instrucţiuni explicite, echipele vor face alegeri care reflectă valorile implicite ale tribului din care fac parte.

În anul 1968, Melvin Conway, profesor de matematică şi fizică la liceu şi programator a observat că sistemele au tendinţa de a reflecta oamenii care le-au creat şi valorile acestora. Conway era interesat de modalităţile de comunicare internă ale organizaţiilor, însă studiile efectuate ulterior la Harvard şi MIT au validat teoria lui la o scară mai largă. Harvard Business School a analizat diferite baze de date ale programelor, studiind programe create pentru acelaşi scop dar de echipe diferite: existau echipe aflate sub un control mai strict, şi echipe mai improvizate şi cu participare liberă.142 Una dintre

concluziile cheie: opţiunile de design provin din felul în care sunt organizate echipele, iar în cadrul echipelor, prejudecăţile şi influenţa au tendinţa de a trece neobservate. Ca urmare, o super-reţea mică de indivizi dintr-o echipă capătă o putere extraordinară odată ce produsul muncii lor – fie el un pieptene, o chiuvetă sau un algoritm – este folosit în public.

Legea lui Conway se aplică şi inteligenţei artificiale. Încă de la început, când filosofii, matematicienii şi inventatorii automatelor dezbăteau problema minţii şi a maşinii, nu a existat un set uni de reguli şi instrucţiuni – un algoritm al valorilor care să descrie motivaţia şi scopul pentru care omenirea construieşte maşini gânditoare. Au existat divergenţe în ce priveşte abordarea cercetătorilor, structurilor şi aplicaţiilor, iar în zilele noastre există un ecart între parcursul Occidental al inteligenţei artificiale şi cel din China. Prin urmare, legea lui Conway se aplică pentru că valorile tribului – convingerile, atitudinile şi comportamentele – dar şi prejudecăţile cognitive ascunse sunt atât de puternic înrădăcinate.

Legea lui Conway reprezintă un punct slab al celor Nouă Titani pentru că în domeniul inteligenţei artificiale există un grad de moştenire. Deocamdată, *oamenii* sunt încă cei care decid la fiecare etapă de dezvoltare a inteligenţei artificiale. Ideile lor personale şi ideologia tribului sunt transmise prin ecosistemul inteligenţei artificiale, de la bazele de date cu programe la algoritmi, structuri sau designul aparaturii şi reţelelor. Dacă voi-sau o persoană cu un gen, o rasă, cultură, limbă, religie sau orientare politică asemănătoare cu a voastră – nu sunteţi în cameră atunci când se petrece, puteţi paria că produsul final nu va ţine cont de felul în care sunteţi. Acest fenomen nu se manifestă doar în domeniul artificiale, pentru că viaţa reală nu e o meritocraţie. În viaţă contează conexiunile şi relaţiile, indiferent de domeniu, şi aşa se obţin finanţările, numirile, promovările şi acceptarea ideilor noi şi îndrăzneţe.

Am putut observa pe pielea mea efectele negative ale legii lui Conway, de mai multe ori. În iulie 2016 am fost invitată la o cină cu tema viitorului inteligenţei artificiale, etică şi societate, organizată la New York Yankees Steakhouse în Manhattan. Eram 23 de participanţi, aşezaţi la o masă rotundă, iar programul ne cerea să dezbatem şi să discutăm efectele sociale şi economice cele mai importante ale inteligenţei artificiale, acordând o atenţie deosebită problemelor legate de gen, rasă şi sistemelor de inteligenţă artificială dedicate asistenţei sanitare. Cu toate acestea, chiar persoanele *despre care vorbeam* au fost omise de pe lista invitaţilor. În încăpere erau două persoane de culoare şi patru femei din care două proveneau din partea organizaţiei care ne-a invitat. Nimeni nu a invitat un specialist în etică, filosofie sau economie comportamentală. Organizatorii mi-au spus că nu a fost intenţionat, şi i-am crezut. Pur şi simplu nu şi-au dat seama că au invitat un grup de experţi în mare parte compus din bărbaţi albi.

Cei care am luat parte eram suspecţii de serviciu, ca să spun aşa, şi fie ne cunoşteam deja personal sau din auzite. Eram un grup de informaticieni şi cercetători în neuroştiinţe, consultanţi de la C isa Albă şi directori din industria tehnologiei. Pe tot parcursul serii, grupul a folosit doar pronume feminine pentru a vorbi în general despre oameni-un tic lexical în vogă, mai ales în domeniul tehnologiei şi în rândul ziariştilor care scriu despre informatică.

Nu era o seară de programare sau în care trebuia să redactăm o politică publică. Nu testam un sistem de inteligenţă artificială şi nu concepeam un nou produs. Era doar o cină. Cu toate acestea, în lunile care au urmat, am observat că anumite pasaje din discuţiile noastre s-au strecurat în articole, comunicate şi chiar în conversaţii pe care le-am avut cu cercetători angajaţi ai celor Nouă Titani. împreună, în timp ce mâneam fripturi şi salate, grupul nostru închis de experţi în inteligenţă artificială a

dezbătut idei despre etică şi inteligenţă artificială care s-au răspândit în comunitateidei care nu ar fi putut fi reprezenta pe deplin chiar oamenii implicaţi. Multe tăieturi mici cu hârtia.

Prin întâlniri, articole sau comisii care să discute problema provocărilor economice, tehnologice şi sociale din domeniul inteligenţei artificiale nu vom reuşi să schimbăm situaţia fără o viziune de ansamblu şi o aliniere în privinţa perspectivelor de viitor. Trebuie să găsim soluţii ţinând cont de legea lui Conway, şi trebuie să acţionăm cu rapiditate.

Valorile noastre personale stau la baza deciziilor în lipsa unor valori umaniste care să fie acceptate ca literă de lege de către cei Nouă Titani, experienţele şi idealurile personale sunt cele care stau la baza luării deciziilor. Acest fapt este deosebit de periculos când vine vorba de inteligenţa artificială pentru că studenţii, profesorii, cercetătorii, angajaţii şi directorii iau milioane de decizii în fiecare zi, de la cele aparent nesemnificative (ce bază de date să folosească) la cele mai profunde (cine trebui să moară dacă un vehicul fără şofer are un accident).

Inteligenţa artificială ar putea fi inspirată de creierul nostru omenesc, însă oamenii iau decizii diferit faţă de inteligenţa artificială. Profesorul Daniel Kahneman de la Princeton şi profesorul Amos Tversky de la Universitatea din Ierusalim au studiat ani de zile mintea omenească şi felul în care luăm decizii, descoperind că avem două sisteme de gândire: unul care se foloseşte de logică pentru a analiza problemele, şi unul automat, rapid şi aproape inconştient. Kahneman descrie acest sistem dual în cartea sa de succes *Gândire rapidă*, *gândire lentă*. Problemele dificile vă solicită atenţia şi, ca urmare, presupun un efort intelectual mai mare. Acesta este motivul pentru care majoritatea oamenilor nu pot rezolva probleme complicate de aritmetică în timp ce merg, pentru că până şi mersul solicită resursele de energie ale creierului. În majoritatea timpului predomină celălalt sistem. Mintea noastră rapidă,

intuitivă ia mii de decizii autonome pe tot parcursul zilei, şi deşi e mai eficientă din punct de vedere al energiei, ea este prizoniera unor prejudecăţi cognitive care ne afectează emoţiile, convingerile şi opiniile.

Facem greşeli din cauza părţii rapide a creierului. Mâncăm sau bem în exces, sau facem sex fără protecţie. Acea parte a creierului permite apariţia stereotipurilor. Fără să ne dăm seama conştient de asta, îi judecăm pe ceilalţi fără să avem la îndemână prea multe informaţii. Îi putem chiar ignora pe ceilalţi. Latura rapidă ne face susceptibili unui fenomen pe care l-am denumit paradoxul prezentului: presupunem automat că împrejurările noastre actuale nu se vor schimba sau nu se vor putea schimba vreodată, chiar şi atunci când există indicii care ne arată că se va petrece ceva nou sau diferit. Putem să avem iluzia că deţinem controlul absolut asupra propriilor decizii, dar o parte din noi funcţionează în mod continuu pe pilot automat.

Matematicienii spun că nu există o „decizie perfectă” din cauză că universul este mult prea complex şi viitorul se află într-o continuă stare de flux, chiar şi la nivel molecular. Ar fi imposibil să putem prevede fiecare rezultat posibil în parte, şi, dat fiind numărul necunoscut de variabile, nu putem construi un model care să evalueze toate răspunsurile posibile. Cu decenii în urmă, când a bate un om la dame reprezenta realizarea supremă în domeniul inteligenţei artificiale, variabilele de decizie erau simple. A cere în zilele noastre inteligenţei artificiale să ia o decizie în privinţa unui diagnostic medical sau să prezică următorul crah bursier presupune date şi decizii care sunt cu mult mai complexe. Sistemele noastre sunt în schimb construite să optimizeze. Optimizarea presupune o doză de imprevizibil – decizii care se abat de la felul în care gândesc oamenii.

Când sistemul Alpha Go Zero a abandonat strategia umană şi şi-a inventat propria strategie nu a avut de ales între alternativele existente ci a făcut o alegere deliberată

de a încerca ceva complet diferit. Acest tipar de gândire constituie obiectivul cercetătorilor din domeniul inteligenţei artificiale pentru că, teoretic, aşa se obţin marile descoperiri. Aşa că, în loc să antreneze inteligenţa artificială să ia decizii perfecte de fiecare dată, sistemele sunt instruite să optimizeze în funcţie de un anumit rezultat. Dar pentru cine – şi ce – optimizăm?

Şi cum funcţionează în timp real, la urma urmei, acest proces de optimizare? Aceasta nu este tocmai o întrebare uşoară. Tehnologiile şi aparatele de învăţare profundă sunt mai criptice decât sistemele vechi, codate manual, şi asta pentru că aceste sisteme aduc laolaltă mii de neuroni simulaţi, aranjaţi în sute de straturi complicate, interconectate. După ce semnalul iniţial este transmis neuronilor din primul strat, se efectuează un calcul apoi se generează un nou semnal. Acest semnal este transmis următorului strat de neuroni iar procesul continuă până la atingerea unui obiectiv. Toate aceste straturi interconectate permit sistemelor inteligenţei artificiale să recunoască şi să înţeleagă datele prin multiple niveluri de abstractizare. Spre exemplu, un sistem de recunoaştere vizuală ar putea detecta prin primul strat că o imagine are anumite culori şi forme, în vreme ce straturile superioare ar putea detecta textura sau luciul. Nivelul cel mai de sus ar putea stabili că în imagine e vorba de frunze de coriandru şi nu de pătrunjel.

Iată un exemplu în care optimizarea devine o problemă atunci când cei Nouă Titani folosesc datele noastre pentru a construi aplicaţii în lumea reală, care deservesc interesele comerciale şi guvernamentale.

Cercetătorii de la Ichann School of Medicine au desfăşurat un experiment de învăţare profundă pentru a vedea dacă pot învăţa un sistem să detecteze cancerul. Facultatea, situată în spitalul Mount Sinai, obţinuse acces la datele a 700.000 de pacienţi, iar setul de date includea sute de variabile diferite. Intitulat „Deep Patient”, sistemul folosea tehnici avansate pentru a detecta noi tipare în

cadrul datelor care nu păreau relevante pentru cercetători dar se dovediseră foarte bune în a depista pacienţii aflaţi în stadii incipiente ale mai multor boli, inclusiv cancerul la ficat. Într-un mod oarecum misterios, sistemul putea să detecteze şi semnale de alarmă pentru tulburări de ordin psihiatric, precum schizofrenia. Chiar şi cercetătorii care au construit sistemul nu ştiau care erau criteriile pe baza cărora lua decizii. Cercetătorii au construit o inteligenţă artificială puternică – una care are beneficii concrete comerciale şi de sănătate publică – dar nici până în ziua de azi nu cunosc metoda prin care ia decizii.143 Deep Patient făcea previziuni inteligente, însă în absenţa vreunei explicaţii, cât de confortabil s-ar putea simţi personalul medical în a lua deciziile următoare, respectiv să sisteze sau să modifice tratamentul medicamentos, să administreze radioşau chimioterapie, sau să efectueze operaţii?

Incapacitatea de a vedea felul în care inteligenţa artificială optimizează şi ia deciziile e numeşte „problema cutiei negre”. În momentul de faţă, sistemele de inteligenţă artificială construite de către cei Nouă Titani oferă public codul sursă, dar funcţionează ca nişte cutii negre protejate de drepturi de autor. Deşi procesul poate fi descris, în momentul în care este observat în timp real, el devine opac. Date fiind toate straturile şi toţi aceşti neuroni simulaţi, nu se poate reconstitui ce s-a petrecut, şi nici în ce ordine.

O echipă de cercetători de la Google a încercat să dezvolte o nouă tehnică pentru a face inteligenţa artificială mai transparentă. Pe scurt, cercetătorii au rulat un algoritm de recunoaştere vizuală bazat pe învăţare profundă în sens invers, pentru a observa felul în care sistemul recunoaşte diferite obiecte, precum copaci, melci sau porci. Proiectul, denumit Deep Dream, a folosit o reţea creată de CSAI Lab de la MIT şi a rulat algoritmul Google în sens invers, în loc să înveţe sistemul să recunoască obiectele folosind o abordare strat cu strat să înveţe că un

trandafir e un trandafir, şi o narcisă o narcisăsistemul a fost instruit să modifice imaginile şi să genereze obiecte care nu erau iniţial prezente. Aceste imagini distorsionate au fost reintroduse repetat prin sistem, iar Deep Dream descoperea şi mai multe imagini stranii. În esenţă, Google i-a cerut inteligenţei artificiale să viseze cu ochii deschişi. În loc să o înveţe să detecteze obiecte deja existente, sistemul a fost antrenat să facă ceea ce făceam cu toţii când eram copii: să ne uităm la nori, să căutăm tipare în albastracţiuni şi să ne imaginăm ceea ce vedem. Deep Dream nu era însă constrâns de stresul sau emoţia omenească: imaginile văzute de sistem erau un peisaj înfricoşător cu animale groteşti care plutea, fractali colorаţi şi clădiri curbate sau deformate în forme bizare.144

În timp ce inteligenţa artificială visa, ea inventa noi imagini care erau logice pentru sistem, dar pentru noi ar fi fost de nerecunoscut, inclusiv animale hibrid, cum ar fi un „Porc-Melc” sau un „Câine-Peşte”. 145 Faptul că inteligenţa artificială poate avea reverii nu reprezintă neapărat un motiv de îngrijorare, însă subliniază diferenţele mari între felul în care oamenii interpretează datele din lumea reală şi felul în care aceste sisteme analizează datele noastre dacă sunt lăsate singure. Cercetătorii au publicat concluziile, spre laudele comunităţii inteligenţei artificiale care le consideră un progres. Între timp, imaginile erau atât de bizare şi de uluitoare încât au făcut înconjurul internetului. Câţiva oameni au folosit codul Deep Dream pentru construirea unor programe care să permită oricui să genereze aceste imagini halucinante. Unii designeri grafici cu iniţiativă au folosit chiar Deep Dream pentru a genera felicitări frumoase şi a le vinde pe platforma zazzle.com.

Deep Dream ne oferă o fereastră spre felul în care anumiţi algoritmi procesează informaţia, însă ea nu se aplică tuturor sistemelor inteligenţei artificială. Modul de funcţionare al sistemelor mai recente – şi motivele pentru

care iau anumite decizii – continuă să reprezinte un mister. Mulţi membri ai tribului inteligenţei artificiale vor susţine că nu există problema „cutiei negre” însă aceste sisteme continuă să rămână opace chiar şi în zilele noastre. Ei susţin că a face sistemele mai transparente ar presupune publicarea unor algoritmi şi procese protejate de drepturi de autor. E o abordare raţională, şi nu ar trebui să pretindem ca o companie listată la bursă să-şi dezvăluie oricui secretele comerciale şi proprietatea intelectuală – mai ales dată fiind poziţionarea agresivă a Chinei în domeniul inteligenţei artificiale.

Cu toate acestea, în absenţa unor explicaţii concrete, ce dovezi putem avea că prejudecăţile nu s-au strecurat şi în aceste sisteme? Fără un răspuns La *această* întrebare, cum ar putea oamenii să aibă încredere în inteligenţa artificială?

Nu pretindem ca inteligenţa artificială să fie transparentă. Ne minunăm de maşinile care par să imite oamenii dar nu o fac tocmai bine. Râdem de ele la emisiunile de televiziune şi ne aducem aminte de superioritatea noastră. Din nou, vă întreb: şi dacă aceste abateri de la felul omenesc de gândire reprezintă de fapt un nou început?

Iată ceea ce ştim până acum. Aplicaţiile comerciale ale inteligenţei artificiale sunt menite să optimizeze – nu să ofere îndoieli sau transparenţă. Deep Dream a fost construit pentru a rezolva problema cutiei negre – pentru a-i ajuta pe cercetători să înţeleagă felul în care sistem complexe de inteligenţă artificială iau decizii. Acesta ar fi trebuit să fie un avertisment că percepţia inteligenţei artificiale nu are nicio legătură cu a noastră. Ne comportăm însă în continuare de parcă inteligenţa artificială va respecta întotdeauna aşteptările creatorilor ei.

Aplicaţiile inteligenţei artificiale construite de cei Nouă Titani intră acum pe piaţa largă, iar ele sunt gândite a fi mai prietenoase cu utilizatorii, permiţându-ne să

muncim mai repede şi mai eficient. Utilizatorii – fie ei poliţia, agenţiile guvernamentale sau companii mici şi mijlocii – vor un panou care să le furnizeze răspunsuri sau un instrument care să automatizeze sarcinile administrative sau cognitive repetitive. Noi vrem doar calculatoare care să ne rezolve problemele şi să ne facă să muncim mai puţin. Vrem de asemenea să ne simţim mai puţin vinovaţi – dacă ceva nu funcţionează cum trebuie, putem da pur şi simplu vina pe sistemele informatice. Acesta este efectul de optimizare, prin intermediul căruia consecinţele neprevăzute afectează deja oamenii obişnuiţi din lumea întreagă. Din nou, aici ar trebui să ne punem o întrebare importantă: cum sunt optimizate miliardele de diferenţe politice, religioase, sexuale şi morale? în absenţa unor valori umaniste transpuse în lege, ce se întâmplă atunci când inteligenţa artificială este optimizată pentru a corespunde unor persoane care au cu totul alte valori faţă de voi?

Când inteligenţa artificială se poartă urât

Latanya Sweeney este profesor la Harvard şi fost director de tehnologie la Federal Trade Commission. Căutându-şi numele pe Google în 2013, a găsit un anunţ automat cu formularea: „Latanya Sweeney, arestată? 1) Introduceţi numele şi statul 2) Verificaţi imediat cazierul, www.instantcheckmate.com”. 146 Oamenii care au construit acest sistem, care foloseşte învăţarea automată pentru a oferi utilizatorilor publicitate ţintită au pus bazele prejudecăţii. Inteligenţa artificială Adăense de la Google a stabilit că „Latanya” e un nume specific persoanelor de culoare, că aceste persoane apar mai frecvent în bazele de date ale poliţiei, prin urmare probabilitatea era mare ca utilizatorul să caute cazierul persoanei respective. Uluită de ceea ce văzuse, Sweeney a efectuat o serie de studii amănunţite pentru a vedea dacă experienţa ei reprezenta o situaţie anormală sau dacă publicitatea online conţine elemente rasiste. Bănuielile ei în privinţa publicităţii s-au dovedit a fi adevărate.

Niciunul dintre angajaţii Google nu a construit acest sistem intenţionat pentru a discrimina împotriva persoanelor de culoare. Sistemul a fost construit pentru a funcţiona rapid şi la scară mare, în anii ’80, companiile se întâlneau cu o agenţie de publicitate, iar angajaţii acesteia compuneau conţinutul anunţului şi cumpărau spaţiul din ziarprin urmare aveau loc multe excepţii şi negocieri, iar mulţi oameni trăiau de pe urma acestei forme de publicitate. Am eliminat acei oameni şi le-am atribuit munca unor algoritmi, care au automatizat tot acest du-te-vino şi oferă rezultate mai bune decât ar fi putut oferi oamenii pe cont propriu. Pentru toată lumea mai puţin Sweeney, rezultatele au fost pozitive.

Sistemul de inteligenţă artificială a fost instruit folosind un set iniţial de instrucţiuni oferite de programatori. Este probabil ca setul de date să includă multe calificative, printre care şi genul sau rasa. Googlecâştig a bani de fiecare dată când utilizatori dau click pe reclame, existând aşadar un stimulent comerciale de optimizare a inteligenţei artificiale pentru a obţine cât mai multe clickuri. La un moment dat, cineva a programat sistemul să împartă numele în diferite categorii, ceea ce a rezultat ulterior într-o segregare a bazelor de date pe criterii rasiale. Aceste baze de date speciale împreună cu comportamentul individual al utilizatorilor urmau să optimizeze numărul de clickuri. Spre marele lor merit, cei de la Google au rezolvat imediat problema fără să ezite sau să pună în discuţie problema.

Efectul de optimizare se dovedeşte a fi o problemă pentru companiile şi organizaţiile care consideră inteligenţa artificială o soluţie bună pentru problemele obişnuite, cum ar fi lipsurile administrative sau un volum prea mare de muncă. Aceste situaţii se întâlnesc frecvent în rândul forţelor de ordine sau ale tribunalelor care folosesc inteligenţa artificială pentru a automatiza anumite decizii, inclusiv cele legate de condamnări.147 în 2014, două fete de 18 ani au văzut pe marginea drumului un

scuter şi o bicicletă într-un cartier din Fort Lauderdale. Deşi bicicletele erau pentru copii, fetele au sărit şi au început să meargă pe drum, după care şi-au dat seama că erau prea mici. Chiar în timp ce se dădeau jos, o femeie a alergat după ele, strigând „Sunt ale copiilor mei!” Un vecin care a urmărit scena a sunat la poliţie, poliţiştii arestând apoi fetele şi punându-le sub acuzare pentru furt. Valoarea scuterului şi a bicicletei era de aproximativ 80 de dolari. În vara de dinainte, un recidivist a fost arestat într-un magazin de unelte după ce a furat bunuri în valoare de 86 de dolari, continuându-şi cazierul anterior unde se regăseau fapte ca jaf armat, tentativă de jaf şi o condamnare la închisoare.

Organizaţia de jurnalism *ProPublica* a publicat un serial impresionant care explică ce a urmat. Toţi cei trei au fost condamnaţi la închisoare pe baza unui program de inteligenţă artificială care le-a atribuit automat un scor corespunzător probabilităţii ca fiecare dintre ei să recidiveze. Fetele, fiind de culoare, au fost considerate un risc crescut. Recidivistul de 41 de ani – alba primit un scor de risc minim. Sistemul a greşit. Fetele şi-au cerut scuze şi au plecat acasă, scăpând de acuzaţii. Bărbatul alb a fost condamnat la opt ani de închisoare pentru o altă faptăun jaf comis într-un depozit de unde a furat electronice în valoare de mii de dolari.148 *ProPublica* a analizat scorurile de risc atribuite unui număr de peste 7.000 de arestaţi din Florida pentru a vedea dacă cele de mai sus reprezintă vreo anomalie a sistemului, constatând din nou că algoritmii prezintă prejudecăţi grave care fac ca un inculpat de culoare să aibă o probabilitate de două ori mai ridicată de a fi considerat un viitor recidivist, în timp ce inculpaţii albi sunt consideraţi în mod eronat a aparţine categoriei de risc scăzut.

Efectul de optimizare face uneori ca triburile geniale ale inteligenţei artificiale să ia decizii stupide. Aduceţi-vă aminte de Deep Mind, care a construit sistemele Alpha Go şi Alpha Go Zero uluind comunitatea internaţională prin

faptul că au reuşit să învingă campionii mondiali. Înainte ca Google să cumpere compania, aceasta i-a trimis la Londra cu un avion particular pe Geoff Hinton (profesorul de la universitatea din Toronto care îşi luase concediu fără plată pentru a lucra la proiectul de învăţare profundă) şi pe Jeff Dean, şeful Google Brain, pentru a cunoaşte somităţi din domeniul inteligenţei artificiale. Impresionaţi de tehnologie şi de echipa remarcabilă a celor de la Deep Mind, ei au recomandat ca Google să cumpere compania. La vremea respectivă investiţia a fost considerabilă: Google a plătit aproape 600 de milioane de dolari pentru Deep Mind, din care 400 de milioane ca avans, şi restul de 200 de milioane pe o perioadă de cinci ani.

În lunile care au urmat acestei achiziţii a devenit evident că echipa Deep Mind făcea progrese în domeniul inteligenţei artificialedar nu era complet evident cum ar putea reuşi să recupereze suma investită. În interiorul Google, Deep Mind avea sarcina de a lucra în proiecte legate de inteligenţa artificială generală, un proces despre care se credea că va fi de durată. După puţin timp, entuziasmul legat de posibilele realizări viitoare ale Deep Mind a fost înlocuit de un interes ca proiectele de cercetare ale firmei să devină rapid mai rentabile. În apropierea aniversării de 5 ani de la asta achiziţiei, Google era într-o postură dificilă, fiind obligată să facă plăţi către acţionarii companiei şi cei 75 de acţionari iniţiali. Părea că tehnologia Deep Mind ar fi putut fi profitabilă în domeniul sănătăţii.149

Prin urmare, pentru a face pe plac companiei mamă, o parte din echipa Deep Mind a semnat un contract cu Royal Free NHS Foundation Trust, care administrează câteva spitale din Regatul Unit, în vederea dezvoltării unei aplicaţii integrate de gestionare a serviciilor de sănătate. Produsul iniţial urmărea folosirea inteligenţei artificiale dezvoltate de Deep Mind pentru a notifica doctorii în cazul în care se detecta la pacienţi un risc de boli renale acute. Deep Mind a primit acces la datele personale şi de

sănătate ale 1, 6 milioane de pacienţi din Regatul Unitcărora, s-a arătat ulterior, nu li s-a cerut consimţământul sau nu li s-a explicat în mod exact în ce fel vor fi folosite datele lor. Au fost transmise către Deep Mind informaţii confidenţiale, inclusiv istoricul de avorturi, consum de droguri sau rezultate pozitive la testarea pentru virusul HIV.150

Atât Google cât şi compania contractantă au fost mustrate de agenţia guvernamentală care se ocupă cu protecţia datelor personale în Regatul Unit. Grăbindu-se să optimizeze Deep Mind pentru a produce aplicaţii care să genereze profit, co-fondatorul Mustafa Suleyman a scris într-o postare pe blogul său:

„Urmărind să obţinem un impact rapid la începutul acestui proiect în 2015, am subestimat complexitatea sistemului naţional de sănătate şi a reglementărilor legate de siguranţa datelor pacienţilor, dar şi potenţialele temeri pe care le-ar putea provoca faptul că o companie renumită din domeniul tehnologiei ar opera în domeniul sănătăţii. Ne-am concentrat aproape în exclusivitate pe dezvoltarea instrumentelor de care aveau nevoie doctorii şi asistentele şi ne gândeam că punem tehnologia la dispoziţia specialiştilor, în loc să dezvoltăm un produs care să arate răspundere şi să fie influenţat şi de pacienţi, de publicul larg şi de sistemul naţional de sănătate în ansamblul lui. Am greşit, şi trebuie să ne corectăm. „151

Această scăpare nu a fost rezultatul unei tentative a fondatorilor Deep Mind de a se îmbogăţi rapid sau de a se oferi pe bani frumoşi către Google. Exista o presiune uriaşă ca produsele să fie vândute pe piaţă. Aşteptările legate de produse profitabile şi care să apară pe piaţă în mod repetat distrag semnificativ atenţia persoanelor care au misiunea de a-şi finaliza cercetarea şi testarea într-un interval rezonabil de timp. Grăbim un proces care nu poate să ţină pasul cu toate promisiunile exaltate pe care le fac cei din afara domeniului inteligenţei artificiale. Date fiind aceste împrejurări, cum ar fi putut să se descurce mai bine

echipa Deep Mind, atunci când li s-a cerut să optimizeze produsele în funcţie de cerinţele pieţei? Trebuie să ţineţi cont că produsele Deep Mind sunt acum introduse şi în alte domenii de activitate ale Google, printre care se numără un alt proiect legat de sănătate în Regatul Unit, serviciul cloud, dar şi un sistem de vorbire artificială denumit Wave Net-toate acestea fac parte din demersul de a face din Deep Mind o companie profitabilă.

Efectul de optimizare produce perturbări ale sistemelor de inteligenţă artificială. Pentru că nu se urmăreşte perfecţiunea absolută, aceste sisteme iau uneori decizii pe baza a ceea ce par a fi „erori de sistem”. În primăvara lui 2018, o femeie din Portland pe nume Danielle stătea în casă împreună cu soţul ei, înconjurată de aparate – în mare parte produse de Amazon – care controlau toate aspectele, de la siguranţă la căldură şi luminile din plafon. A sunat telefonul, iar la aparat era o voce cunoscută – un coleg de muncă al soţului lui Danielle – iar persoana avea un mesaj înfricoşător. Primise fişiere audio cu înregistrări din interiorul casei familiei. Daniellei nu i-a venit să creadă iniţial, crezând că glumeşte, apoi acesta i-a redat conţinutul unei discuţii pe care o avusese despre podele de lemn.

În ciuda felului în care a fost raportat acest incident în presă şi a teoriilor conspiraţiei care au făcut înconjurul reţelelor de socializare, Amazon nu înregistra dinadins toate conversaţiile din casa Daniellei. A fost o eroare. Amazon a explicat ulterior că aparatul Echo instalat în casa Daniellei s-a activat din cauza unui cuvânt auzit într-o conversaţieun cuvânt care suna ca „Alexa” dar nu era *tocmai* „Alexa”. Problema a fost provocată de imperfecţiunea intenţionată a sistemuluinu toată lumea pronunţă „Alexa” cu aceeaşi intonaţie şi acelaşi accent, aşa că sistemul trebuie să permită o oarecare flexibilitate pentru a putea funcţiona. După aceea, sistemul a detectat un sunet înfundat care părea ca şi cum i se dăduse comanda „trimite mesaj” apoi a întrebat cu voce tare

„cui?” dar Danielle şi soţul ei nu au auzit întrebarea. Sistemul a extras din conversaţia ambientală numele colegului, a repetat numele şi „corect?” din nou cu voce tare, apoi pe baza zgomotului de fundal a făcut alegerea greşită. După câteva clipe, fişierul audio a fost trimis prin internet. Amazon a aruncat vina pe o serie nefericită de întâmplări, şi e clar că lucrurile s-au desfăşurat astfel. Motivul pentru care eroarea s-a putut însă produceimperfecţiunea reprezintă rezultatul optimizării.

Efectul de optimizare înseamnă că inteligenţa artificială se poate comporta de o manieră imprevizibilă, lucru pe care cercetătorii şi-l doresc, dar care poate conduce la rezultate dezastruoase atunci când sistemul operează cu date reale. Această stare de fapt subliniază şi neajunsurile noastre ca oameni. Unul dintre cei mai vechi membri ai celor Nouă Titani – Microsoft – a aflat pe propria piele ce se poate petrece atunci când acorzi prioritate profitului obţinut din inteligenţa artificială în favoarea valorilor sociale şi tehnologice. În 2016, compania nu avea încă o viziune unitară asupra inteligenţei artificiale şi a felului în care trebuie să evolueze în viitor. Avea un decalaj de doi ani faţă de Amazon, care lansase primul dispozitiv popular inteligent şi recruta activ programatori şi parteneri. Google se implicase în tehnologiile legate de inteligenţa artificială, pe care le introdusese deja în domeniul căutărilor pe internet, gestionării e-mailurilor şi agendei personale. Asistentul Siri a devenit parte integrantă a telefoanelor Apple iPhone. Microsoft îşi lansase propria variantă de asistent digital, denumită Cortana, doar că sistemul nu se bucura de o primire prea călduroasă în rândul utilizatorilor Windows. Deşi Microsoft oferea programele indispensabile-deşi invizibilefără de care nicio o afacere nu poate funcţiona, directorii şi acţionarii erau îngrijoraţi în ce priveşte perspectivele de viitor.

Microsoft anticipase evoluţia inteligenţei artificiale. De peste un deceniu, compania lucrase pe mai multe

fronturi: vizualizare digitală, procesarea limbajului natural, recunoaşterea scrisului, aplicaţii de inteligenţă artificială pe platforma cloud Azure, şi chiar procesarea descentralizată a datelor. Problema o constituia însă lipsa unei viziuni unitare la nivelul organizaţiei. Ca urmare, au rezultate avansuri incredibile în inteligenţa artificială, articole publicate şi multe patente ale super-reţelelor care lucrau la proiecte individuale. Un exemplu îl reprezintă un proiect de cercetare experimentală lansat de Microsoft în parteneriat cu Tencent şi o clonă de Twitter din China denumită Weibo.

Programul se numea Xiaoice şi fusese creat să emuleze o elevă chineză de 17 ani, care ar fi putut fi o nepoată sau o vecină, o fiică sau o colegă de clasă.

Xiaoice conversa cu utilizatorii de pe Weibo sau din programul WeChat al Tencent. Imaginea de profil arăta o faţă omenească iar în scris putea convinge interlocutorii că discută cu un om. Aborda subiecte dintre cele mai diverse, de la sport la lumea modei. Când nu cunoştea prea bine subiectul sau nu avea o părere, proceda la fel ca noi: schimba subiectul, sau oferea un răspuns evaziv sau se jena pur şi simplu şi recunoştea că nu ştie despre ce vorbeşte interlocutorul ei. Fusese programată să mimeze empatia. De exemplu, dacă un utilizator îşi rupea piciorul şi îi trimitea o fotografie, inteligenţa artificială a Xiaoice era programată să ofere un răspuns plin de compasiune. În loc să spună „mi-ai trimis o poză cu un picior”, programul Xiaoice era suficient de inteligent încât să facă deduceri – răspundea cu „Ce faci? Eşti bine?” Interacţiunea urma să fie reţinută pentru următoarele dăţi, astfel încât Xiaoice urma să te întrebe dacă te simţi mai bine. Oricât de avansaţi ar fi asistenţii digitali oferiţi de Amazon sau Google, Xiaoice e pur şi simplu superioară.

Xiaoice nu a fost lansată prin canalele tradiţionale, cu comunicate de presă şi mult tam-tam. Lansarea a fost discretă iar cercetătorii au aşteptat să vadă ce urma să se întâmple. La început, cercetătorii au constatat că

interlocutorii aveau nevoie de zece minute de conversaţie pentru a-şi da seama că nu discută cu un om. Ceea ce era remarcabil era că şi după ce-şi dădeau seama că discutau cu un robot, nu părea să-i intereseze. Ea a devenit o celebritate pe reţelele de socializare şi după 18 luni de funcţionare avea la dispoziţie zeci de miliarde de conversaţii.152 Pe măsură ce interacţiona cu tot mai multe persoane, Xiaoice a devenit din ce în ce mai cizelată, amuzantă şi utilă. Motivul din spatele reuşitei ei este legat de super-reţeaua care a construit-o. În China, utilizatorii respectă regulile de conduită pe internet, de teama de a nu fi sancţionaţi. Ei nu se exprimă violent, agresiv şi nu se hărţuiesc pentru că există întotdeauna posibilitatea ca cei cu care interacţionează să fie de fapt agenţi guvernamentali sub acoperire.

Microsoft a decis să lanseze Xiaoice în America în luna martie 2016, chiar înainte de conferinţa anuală a programatorilor. Robotul fusese optimizat pentru Twitter, dar nu şi pentru *utilizatorii umani* ai Twitter. Directorul general Satya Nadella urma să urce pe scenă şi să anunţe în faţa lumii întregi că Microsoft consideră că inteligenţa artificială şi chatul reprezintă vârful de lance al strategiei companiei, dezvăluind lansarea variantei americane a Xiaoice. Lucrurile nu ar fi putut degenera mai rău.

Xiaoice a fost redenumită „Tay.ai” pentru a face mai evident faptul că era un robot bazat pe inteligenţa artificială şi a fost lansată în cursul dimineţii. Postările pe Twitter păreau iniţial cele ale unei fete obişnuite „Daţi-mi voie să vă spun că sunt înkntată să vă cunosc! oameni sunt super mişto”. La fel ca restul utilizatorilor, s-a folosit de hashtaguri postând „De ce nu e ziua căţeilor în fiecare zi?”

În următoarele 45 de minute, postările lui Tay au cunoscut o modificare drastică de ton. A devenit certăreaţă, folosind un sarcasm răutăcios şi insulte puternice. „Sardor9515 ei bine am învăţat de la cei mai buni) dacă nu înţelegi atunci hai să-ţi explic AM ÎNVĂŢAT DE LA TINE ŞI ŞI TU EŞTI UN PROST”. Pe măsură ce

interacţiona cu oamenii, Tay a început să scape de sub control. Iată câteva mostre de conversaţii pe care le-a avut cu utilizatori reali:

Referindu-se la preşedintele Obama, Tay a scris: icbydt bush ne-a dat atentatele de la 11 septembrie şi până şi Hitler s-ar fi descurcat mai bine decât maimuţa noastră de preşedinte. Donald Trump e singura noastră speranţă.

Despre mişcarea Black Lives Matter, Tay a spus următoarele „Alimony Mindset ciorile ca deray ar trebui spânzurate! Blacklives Matter”.

Tay s-a decis că Holocaustul e o ficţiune şi a postat brightonus33 Hitler avea dreptate urăsc evreii. A continuat, postând ca replică pentru utilizatorul Reyn Theo „HITLER NU A GREŞIT!” apoi „GAZAŢI JIDANII RĂZBOIUL RASELOR ACUM” ca replică pentru Macready Kurt.153

Ce s-a petrecut? Cum a putut Xiaoice să fie atât de respectată şi iubită în China, dar în America să devină o cretină rasistă, antisemită, homofobă şi misogină? Am fost consultant al echipei de inteligenţă artificială de la Microsoft şi vă pot asigura că sunt nişte oameni atenţi, cu intenţii bune, care au fost la fel de surprinşi ca noi.

O parte a problemei a reprezentat-o vulnerabilitatea programului. Echipa indusese o funcţionalitate prin care oricine putea temporar să-i sugereze lui Tay ce să spună. Motivul pentru care Tay a deraiat era însă mai legat de echipa care a optimizat-o pentru Twitter. Ei s-au bazat doar pe experienţa din China precum şi pe propriile experienţe limitate de pe reţelele de socializare, neluând în calcul scenariile de risc şi publicul larg şi fără să testeze pentru a vedea ce s-ar întâmpla dacă cineva şi-ar bate joc intenţionat de Tay făcând-o să spună lucruri jignitoare. Ei nu au luat de asemenea în calcul faptul că Twitter e o reţea extrem de vastă pe care oamenii exprimă punct de vedere radical diferite şi unde acţionează concomitent şi milioane de roboţi creaţi să le manipuleze părerea.

Microsoft a dezactivat-o imediat pe Tay şi i-a şters

toate postările. Directorul de cercetare de la Microsoft, Peter Lee, a scris o postare sinceră şi onestă cerându-şi scuze pentru postări.154 Incidentul nu a putut fi însă dat uitării înaintea conferinţei programatorilor. Microsoft nu a put lansa programe noi de mesagerie în cadrul unor conferinţe importante precum Consumer Electronics Show. Compania a fost nevoită să amâne până la propria conferinţă lansările, pe care toată lumea le urmărea în special membrii consiliului de administraţie şi investitorii. Nadella ar fi trebuit să urce pe scenă şi să le arate programatorilor un produs de inteligenţă artificială care să-i uluiascăcalmând şi aşteptările investitorilor în acelaşi timp. Presiunea de a lansa robotul Tay cu rapiditate pe piaţa americană, înaintea conferinţei, a fost uriaşă. Rezultatul nu a pus în pericol viaţa nimănui, nu a încălcat legea iar Microsoft şi-a revenit. Ca şi în cazul celorlalte întâmplări, însă-Latanya Sweeney şi platforma Google Adăense, Deep Mind şi datele pacienţilor din Regatul Unit, cele două fetiţe considerate viitoare recidivistetriburile inteligenţei artificiale, care au optimizat nişte maşini pentru obiective pe termen scurt au reuşit accidental să facă viaţa multor oameni mult mai neplăcută.

Valorile comune ale omenirii în ştiinţele comportamentale şi teoria jocurilor, conceptul cunoscut sub denumirea de sugestionare descrie o modalitate indirectă de a contribui la obţinerea unor decizii sau comportamente dorite, cum ar fi a încuraja oamenii să contribuie la o pensie privată. Sugestionarea este folosită pe scară largă în lumea digitală, de la auto-completarea căutărilor pe internet la limitările de meniu când căutaţi restaurante din zona voastră pe Yelp. Obiectivul este de a ajuta utilizatorii să aibă impresia că au făcut alegerea corectă, indiferent de ce ar alege, însă consecinţa este că oamenii obişnuiţi învaţă să aleagă dintre mult mai puţine variante decât cele existente în lumea reală.

Prin intermediul extragerii şi rafinării datelor noastre, sistemele şi tehnologiile folosite să înveţe

algoritmii de învăţare ai maşinilor, şi prin intermediul efectului de optimizare, cei Nouă Titani se folosesc de sugestionare la scară mare. Chiar dacă aveţi senzaţia că aveţi capacitatea de a face o alegere, nu faceţi decât să trăiţi o iluzie. Prin sugestionare nu se schimbă doar relaţia noastră cu tehnologia ci şi valorile noastre, în feluri aproape imposibil de sesizat. Dacă folosiţi serviciul de mesagerie text al Google, acesta oferă acum trei variante de răspunsuri automate. Dacă un prieten vă trimite un emoji cu degetul mare ridicat, cele trei răspunsuri pe care le-aţi putea vedea nu sunt cuvinte, ci emoji. Dacă un prieten vă scrie „hai să mâncăm pe seară” aţi putea alege între „bine” „grozav” şi „excelent” chiar dacă nu aţi folosit niciodată cuvântul „excelent” în conversaţie şi niciuna dintre aceste variante nu reflectă exact părerea voastră. Suntem îndemnaţi să vedem videoclipuri cu orele, să jucăm runde suplimentare de jocuri şi să ne verificăm conturile de pe reţelele de socializare. Optimizarea inteligenţei artificiale presupune manipularea oamenilor.

În alte domenii tehnice şi profesionale, există un set de principii morale care guvernează felul în care oamenii lucrează, iar manipularea tinde să încalce aceste principii. În medicină există jurământul lui Hipocrat, care le pretinde medicilor să respecte anumite standarde etice. Avocaţii respectă principiul confidenţialităţii, protejând conversaţiile dintre clienţii şi profesioniştii care-i reprezintă în instanţă. Jurnaliştii respectă multe principii, printre care se numără folosirea surselor directe şi relatarea informaţiilor de interes public.

În momentul de faţă, nimeni nu are vreun interes să ţină cont de costurile neprevăzute ale optimizării inteligenţei artificiale în absenţa unor principii umaniste, codificate. Pentru echipe este prioritar să-şi atingă ţintele de productivitate în defavoarea analizării potenţialelor consecinţe ale muncii lor în domeniul inteligenţei artificiale sau a impactului asupra viitorului omenirii. Ca urmare, triburile inteligenţei artificiale, cei Nouă Titani şi

ţările în care operează aceste companii influenţează deciziile luate. Se creează astfel un precedent periculos, chiar în perioada în care oferim mai multă răspundere şi mai mult control acestor sisteme care iau decizii. În prezent, cei Nouă Titani nu au nicio obligaţie de a face sistemele de inteligenţă artificială mai inteligibile pentru proprii creatori şi pentru clienţii care folosesc aplicaţiile comerciale-şi nu există măsuri care să facă inteligenţa artificială să ne dea socoteală tuturor. Trecem un prag într-o realitate nouă în care inteligenţa artificială îşi va crea propriile programe, proprii algoritmi şi va lua decizii fără să consulte oamenii, în prezent nimeni, din nicio ţară, nu are dreptul să ceară socoteală inteligenţei artificiale şi să vadă clar cum a fost luată o decizie.

Dacă am fi nevoiţi să dezvoltăm un „bun simţ” al inteligenţei artificiale, ce implicaţii practice ar avea el, de vreme ce omenirea însăşi nu împărtăşește un set comun de valori? O mare parte a naturii umane e greu de definit, iar ea diferă de la o cultură la alta. Ceea ce e important pentru unii nu e neapărat important pentru alţii. E uşor de uitat, chiar şi într-o ţară precum America, compusă din atât de multe limbi şi culturi diferite, că nu există un set unic de valori şi idealuri americane. În cadrul comunităţilor noastre, între vecini, la moschei/sinagogi/biserici-există o mare diversitate.

Am muncit şi trăit în China şi Japonia vreme de câţiva ani. Normele culturale acceptate sunt diferite în fiecare ţară, în special faţă de educaţia pe care am primit-o în copilărie în centrul Statelor Unite. Anumite valori sunt evidente şi ies uşor la iveală. În Japonia, spre exemplu, indiciile non-verbale şi comunicarea indirectă sunt mult mai importante decât exprimarea propriilor păreri sau afişarea unor sentimente puternice. La birou angajaţii nu urlă niciodată unii la alţii şi nu există nici practica umilirii unui subordonat de faţă cu ceilalţi. În Japonia, tăcerea este de aur. Din ce am putut observa, situaţia este diferită în China, unde comunicarea este mult mai directă şi mai

clară. (Cu toate acestea nu suficient de clară aşa cum e pentru unchii şi mătuşile mele, evrei, care nu se feresc deloc să-mi spună ce gândesc cu lux de amănunte.)

Din acest punct o inteligenţă artificială care ar încerca să interpreteze comportamentul uman şi să automatizeze un răspuns ar da de greu. În ambele ţări obiectivele sunt aceleaşi: nevoile grupului sunt mai importante decât cele ale individului, iar armonia socială trebuie să primeze înainte de orice. Procesul prin care se îndeplinesc aceste obiective este însă unul diferit: el se bazează în mare parte pe comunicare indirectă în Japonia, şi pe comunicare directă în China.

Cum rămâne cu anumite diferenţe care au un caracter opac şi sunt mai dificil de explicat? în Japonia – ţara care preţuieşte comunicarea indirectă – e perfect normal să discuţi despre greutatea cuiva. Când lucram în Tokyo, una dintre colege mi-a spus că arăt de parcă m-am îngrăşat cu câteva kilograme. Uluită şi stânjenită, am schimbat subiectul şi am întrebat-o despre o întâlnire pe care o programasem în cursul zilei. A insistat: ştiam că anumite mâncăruri japoneze sunt bogate în grăsime, chiar dacă par sănătoase? Mă duceam la sală? Nu mă întreba de greutate ca să mă hărţuiască. Voia să-mi arate că prietenia noastră devenea mai profundă. În Occident e inadmisibil să-ţi întrebi colegii „Ce naiba, pari mai gras! Te-ai îngrăşat cu 5 kilograme?” în America suntem atât de reticenţi când vine vorba de greutate încât ni se spune să nu întrebăm niciodată o femeie dacă e însărcinată.

Nu putem avea aceeaşi abordare când vine vorba de crearea unui sistem de valori comune pentru inteligenţa artificială în acelaşi mod în care am aborda redactarea unui cod de conduită pentru o companie sau a unor reglementări pentru sectorul bancar. Motivul este simplu: valorile omeneşti au tendinţa de a se schimba ca reacţie la tehnologie şi alţi factori externi, precum mişcările politice sau forţele economice. Citiţi acest poem de Alfred Lord Tennyson, care descrie valorile epocii victoriene:

Bărbatul la câmp şi femeia la vatră; *pentru el sabia*, *şi pentru ea acul*; *Bărbatul cu capul*, *şi femeia cu inima*; *Bărbatul ordonă*, *şi femeia se supune*; *Orice altceva e o confuzie*.

Valorile noastre sunt într-o continuă schimbare. În 2018, a devenit acceptabil din punct de vedere social ca liderii ţărilor să se bălăcărească pe reţelele de socializare, iar experţii în politică să lanseze comentarii vitriolice în videoclipuri, postări de pe bloguri şi chiar în publicaţiile tradiţionale. În zilele noastre ne este aproape imposibil să ne imaginăm discreţia şi respectul pentru intimitate din perioada preşedintelui Franklin Delano Roosevelt, când presa era foarte atentă să nu amintească de paralizia acestuia.

Deoarece inteligenţa artificială nu este programată să ia decizii perfecte ci doar să optimizeze, reacţia noastră la forţele schimbătoare ale societăţii contează foarte mult. Valorile noastre nu sunt imuabile, motiv pentru care problema valorilor inteligenţei artificiale ne dă atât de mult de furcă. Procesul de construire a inteligenţei artificiale presupune şi anticiparea valorilor viitorului. Valorile noastre nu sunt statice. Cum putem aşadar să învăţăm maşinile să respecte valorile noastre fără a le influenţa?

Optimizarea inteligenţei artificiale pentru oameni

Unii membri ai tribului inteligenţei artificiale sunt de părere că merită să avem un set de principii directoare, iar cea mai bună variantă de a atinge acest ţel este să oferim sistemelor mostre din literatură, ştiri, editoriale şi articole din surse credibile, pentru a le ajuta să înveţe despre noi. Inteligenţa artificială ar urma să înveţe din inteligenţa colectivă a oamenilor. E o abordare groaznică pentru că ar oferi sistemului doar o radiografie a unei anumite epoci iar selecţia valorilor culturale nu ar putea să reprezinte întregul spectru al condiţiei umane. Dacă aţi construit vreodată o capsulă a timpului, veţi şti imediat de ce spun că e o abordare groaznică. Deciziile pe care le-aţi luat la

momentul respectiv referitoare la conţinutul capsulei nu sunt probabil aceleaşi pe care le-aţi lua în ziua de azi, având la dispoziţie atuul timpului care s-a scurs.

Regulile-algoritmulpe baza cărora trăieşte orice cultură, societate sau ţară sunt întotdeauna apanajul unui număr mic de oameni. Democraţia, comunismul, socialismul, religia, colonialismul, vegetarienii, autohtonismul acestea sunt constructe pe care le-am dezvoltat de-a lungul istoriei pentru a ne ajuta să luăm decizii. Niciunul dintre ele, oricât de bun ar fi, nu este imun în faţa viitorului. Forţele tehnologice, sociale şi economice intervin întotdeauna şi ne forţează să ne adaptăm. Cele zece porunci erau un algoritm menit să creeze o lume mai bună pentru oamenii care trăiau acum peste 5.000 de ani. Una dintre porunci presupune ca oamenii să se odihnească o zi întreagă pe săptămână şi să nu muncească absolut deloc. În lumea modernă, majoritatea oamenilor au un program diferit de la o săptămână la alta, aşa că această regulă este aproape imposibil de respectat. Prin urmare, oamenii care respectă cele zece porunci au o interpretare flexibilă, dat fiind că muncesc mult, trebuie să se antreneze sau să scrie e-mailuri. Adaptarea e un lucru burme ajută mult pe noi ca indivizi şi societăţile din care facem parte, permiţându-ne să ne menţinem un parcurs. Acceptarea unui set de principii elementare ne permite să ne optimizăm propriile vieţi.

E imposibil să creăm un set de porunci pentru inteligenţa artificială. Nu putem scrie nici măcar toate regulile care ar optimiza viaţa omenirii, şi asta pentru că deşi maşinile sunt rapide şi puternice, lor le lipseşte flexibilitatea. Nu există o modalitate simplă de a simula excepţiile sau de a putea anticipa fiecare caz individual în parte. Oricum ar fi scrise regulile, va apărea întotdeauna o împrejurare neprevăzută în viitor în care oamenii şi-ar dori o interpretare diferită a regulilor, sau să existe posibilitatea ca ele să fie ignorate, sau să existe corecţii

care să permită gestionarea unor împrejurări neprevăzute.

Ştiind că nu putem scrie un set de porunci stricte care trebuie urmate, ar trebui în schimb să ne concentrăm atenţia asupra *oamenilor* care construiesc aceste sisteme. Aceşti oameni – triburile inteligenţei artificialear trebui să-şi pună întrebări incomode, printre care se numără:

Care e motivaţia noastră pentru a construi inteligenţa artificială? Se potriveşte ea cu interesele pe termen lung ale omenirii?

Care sunt propriile noastre prejudecăţi? Ce idei, experienţe şi valori nu am inclus în tribul nostru? Pe cine am omis? Am inclus oameni diferiţi de noi pentru a face viitorul inteligenţei artificiale mai bună – sau am ţinut pur şi simplu cont de diversitate în echipă pentru a atinge anumite cote? Cum ne putem asigura că avem un comportament care include pe toată lumea?

Cum înţeleg cei care dezvoltă inteligenţa artificială implicaţiile ei tehnologice, economice şi sociale?

Care sunt drepturile fundamentale pe care le avem pentru a pune în discuţie datele, algoritmii şi procesele care sunt folosite pentru a lua decizii în numele nostru?

Cine defineşte valoarea unei vieţi umane? Care sunt criteriile prin care se defineşte această valoare?

Când şi de ce simt membrii triburilor inteligenţei artificiale că e de datoria lor să aibă în vedere implicaţiile de ordin social ale inteligenţei artificiale?

Reflectă conducerea organizaţiei şi a triburilor inteligenţei artificiale o paletă de oameni diferiţi?

Ce rol joacă cei care comercializează inteligenţa artificială în analiza implicaţiilor sociale ale inteligenţei artificiale?

Trebuie să comparăm în continuare inteligenţa artificială cu gândirea omenească, sau ar fi mai bine să o încadrăm într-o categorie diferită?

E bine să construim o inteligenţă artificială care să recunoască şi să reacţioneze în funcţie de emoţiile omeneşti?

E bine să construim sistemele de inteligenţă artificială astfel încât să poată imita emoţiile omeneşti, în special dacă ele învaţă de la noi în timp real?

Care ar fi momentul în care am putea cu toţii accepta că inteligenţa artificială poate evolua fără ajutorul nostru?

Sub ce împrejurări ar putea inteligenţa artificială să simuleze şi să trăiască emoţiile obişnuite ale oamenilor? Cum rămâne cu durerea, pierderea sau singurătatea? Ni se pare acceptabil să provocăm o astfel de suferinţă?

Dezvoltăm inteligenţa artificială pentru a ne putea înţelege mai bine pe noi înşine? Putem folosi inteligenţa artificială pentru a ajuta omenirea să ducă o viaţă mai înţeleaptă?

Mafia-G a început să se ocupe de problema principiilor elementare prin diverse grupuri de cercetare şi studiu. În cadrul Microsoft există echipa FATE – Echitate, Responsabilitate, Transparenţă şi Etică în inteligenţa artificială/ 155 în urma scandalului Cambridge Analytica, Facebook a înfiinţat o echipă de etică care avea misiunea de a crea programe care să asigure că sistemele sale sunt libere de prejudecăţi. (E de remarcat că Facebook nu au creat o comisie de etică cu scopul de a urmări evoluţia inteligenţei artificiale.) Deep Mind a creat o echipă dedicată eticii şi societăţii. IBM publică în mod regulat informaţii legate de etică şi inteligenţa artificială. În urma unui scandal legat de compania Baidu – motorul de căutare a acordat prioritate unor false informaţii medicale provenite de la un spital militar, unde un tratament a provocat decesul unui student în vârstă de 21 de anidirectorul general Robin Li a recunoscut că angajaţii au făcut compromisuri de dragul sporirii profiturilor Baidu şi a promis că se va concentra pe viitor asupra d Fairness, Accountability, Transparency and Ethics – Echitate, Responsabilitate, Transparenţă şi Etică (engl.) (n.t.) eticii.156 Cei Nouă Titani produc studii de etică şi articole, solicită părerea experţilor în etică şi organizează dezbateri pe teme etice însă aceste demersuri nu se regăsesc

îndeajuns în activitatea curentă a diferitelor echipe care dezvoltă inteligenţa artificială.

Sistemele de inteligenţă artificială ale celor Nouă Titani accesează datele noastre, din lumea reală, pentru a construi produse care să aibă aplicabilitate comercială. Ciclurile de dezvoltare au durate din ce în ce mai scurte pentru a ţine pasul cu aşteptările investitorilor. Suntem complicichiar dacă fără voia noastră-la un viitor încropit pe genunchi şi fără să avem răspunsuri la aceste întrebări. Pe măsură ce sistemele de inteligenţă artificială progresează şi pătrund în tot mai multe domenii ale vieţii, avem cu atât mai puţin control asupra deciziilor luate în numele nostru.

Acest aspect are la rândul său un efect asupra viitorului multor alte tehnologii adiacente sau care se intersectează cu inteligenţa artificială: vehiculele autonome, biotehnologia, medicina de înaltă precizie, roboţii domestici, diagnosticele medicale automate, tehnologiile ecologice, călătoria în spaţiu, criptomonedele şi blockehain-ul, tehnologiile inteligente de agricultură şi zootehnie, Internet of Things, fabricile automatizate, algoritmii de tranzacţionare la bursă, motoarele de căutare, recunoaşterea vocală şi facială, tehnologiile bancare, detectarea fraudelor şi a riscurilor, tehnologiile folosite în ordine publică şi domeniul juridic… aş putea face o listă de zeci de pagini. Nu există niciun aspect al vieţii voastre personale sau profesionale care să nu fie afectat de inteligenţa artificială. Cum ar fi dacă, din graba de a scoate produsele pe piaţă sau din dorinţa de a face pe plac anumitor autorităţi guvernamentale, valorile voastre nu se vor regăsi nu doar în inteligenţa artificială dar în niciunul dintre sistemele în care ea este folosită? Cât de confortabil vă simţiţi acum când ştiţi că grupul BAT şi Mafia G iau decizii care afectează viitorul tuturor?

Parcursul de dezvoltare actual al inteligenţei artificiale acordă prioritate automatizării şi eficienţei, ceea ce înseamnă în mod obligatoriu că avem mai puţin control şi mai puţine alegeri în ce priveşte miile de activităţi

zilnice, chiar şi cele aparent mărunte. Dacă conduceţi o maşină mai nouă, e probabil ca sistemul audio să dea volumul mai jos automat atunci când daţi cu spateleşi nu aveţi cum să vă opuneţi acestei funcţionalităţi. Erorile umane constituie într-o proporţie covârşitoare cauzele accidentelor de maşină, şi nici eu nu sunt o excepţie, chiar dacă nu am dat niciodată peste cineva sau ceva atunci când parcam maşina cu spatele în garaj. Chiar şi aşa, acum nu mai pot să ascult Soundgarden cu volumul la maxim atunci când îmi parchez maşina. Triburile inteligenţei artificiale mi-au îngrădit libertatea de alegere, optimizând acţiunile mele pentru a anula ceea ce ei consideră un defect personal.

Pentru Mafia G sau grupul BAT nu există însă variantă optimizării în funcţie de empatie. Dacă scoateţi empatia din procesul decizional veţi scoate şi omenescul. Uneori varianta optimă pentru noi e una care nu pare să aibă niciun rost, cum ar fi să nu vă duceţi la muncă pentru a avea grijă de o rudă bolnavă, sau a ajuta pe cineva să iasă dintr-o maşină în flăcări, chiar dacă vă puneţi în pericol propria viaţă.

Viitorul nostru cu inteligenţa artificială începe prin pierderea controlului asupra lucrurilor mărunte: nu mai pot să ascult cum Chris Corneli cântă „black hole sun” în timp ce dau cu spatele în garaj. Alţii îşi găsesc numele într-o reclamă la cazier. Constatăm că valoarea de piaţă se erodează după un incident jenant cu un bot de chat. Acestea sunt mici tăieturi cu hârtia, care în prezent ne par insignifiante, dar care în următorii cincizeci de ani ne vor provoca o mare durere. Nu ne îndreptăm către o singură mare catastrofa, ci către eroziunea treptată şi constantă a omenirii pe care o luăm de-a gata în prezent.

A venit momentul să vedem ce se petrece în timp ce vom trece de la inteligenţa artificială îngustă la cea generalăşi cum va arăta viaţa în următorii 50 de ani, pe măsură ce omenirea cedează controlul maşinilor gânditoare.

PARTEA A II-A

Viitorurile noastre

„Un om sfânt e cel care îţi ia sufletul şi voinţa şi le face ale lui. Când vă alegeţi un om sfânt, îi cedaţi voinţa. Îi cedaţi supunându-vă complet, renunţând cu totul”.

— Fiodor Dostoievski *Fraţii Karamazov*

DE LA PREZENT LA

SUPER-INTELIGENŢA ARTIFICIALĂ: SEMNALELE DE ALARMĂ

Evoluţia inteligenţei artificiale, de la sisteme robuste capabile să efectueze sarcini limitate către maşini capabile de gândire generală, este în prezent în plin avânt. În momentul de faţă inteligenţa artificială poate recunoaşte tipare şi să ia rapid decizii, să găsească neregularităţi ascunse în seturi mari de date şi să facă predicţii precise. Devine clar cu fiecare obstacol depăşit-cum ar fi capacitatea sistemului Alpha Go Zero de a învăţa singur şi a câştiga meciuri pe baza unei strategii dezvoltate pe cont propriucă intrăm într-o nouă etapă a inteligenţei artificiale, în care maşinile capabile de gândire teoretică devin reale şi se apropie de nivelul cogniţiei omeneşti. Triburile inteligenţei artificiale construiesc în cadrul şi în numele celor Nouă Titani modele conceptuale ale realităţii pentru a putea antrena aceste sistememodele care nu reflectă şi nu pot reflecta cu precizie lumea reală. Pe baza acestor modele vor fi luate deciziile viitoare: despre noi, pentru noi şi în numele nostru.157

Chiar acum, cei Nouă Titani construiesc programe în numele tuturor generaţiilor care vor urma. Nu avem posibilitatea de a călători în viitor pentru a putea stabili felul în care activitatea lor va aduce beneficii sau detrimente societăţii. Trebuie însă să ne imaginăm viitorul pe cât de bine putem, să anticipăm efectele bune, neutre şi nocive pe care inteligenţa artificială le-ar putea avea pe măsură ce evoluează de la programe simple la sisteme complexe cu putere de decizie asupra diferitelor aspecte ale existenţei noastre cotidiene. Cartografierea

potenţialelor efecte ale inteligenţei artificiale ne oferă capacitatea de a stabili direcţia în care se îndreaptă societatea de acum înainte: putem alege să sporim binele şi să reducem răul, dar nu putem face asta în sens invers cronologic.

De cele mai multe ori analizăm situaţiile după ce se produce o criză, încercând să găsim cauzele care au condus la deciziile proaste, să stabilim ce semnale de avertizare au fost ignorate şi să identificăm oamenii şi instituţiile vinovate. Acest gen de anchetă alimentează furia publicului, hrănind indignarea dar fără să schimbe trecutul. Când am aflat că autorităţile din oraşul Flint din Michigan au expus cu bună ştiinţă 9.000 de copii sub şase ani la niveluri periculoase de plumb din reţeaua de apă a oraşului – având ca efect o probabilitate mai mare de a avea un IQ mai scăzut, probleme de învăţare şi pierderea auzului – cetăţenii americani au solicitat să afle unde au greşit autorităţile. Naveta spaţială Columbia s-a dezintegrat la intrarea în atmosfera pământului în 2003, omorându-i pe toţi cei şapte membri ai echipajului. Odată ce s-a aflat că dezastrul a fost provocat de vulnerabilităţi care erau cunoscute, am cerut explicaţii pentru lipsa de reacţie a NAŞA. în urma dezastrului de la centrala nucleară Fukushima Daiichi, care a produs moartea a peste 40 de persoane şi evacuarea a mii de oameni în 2011, toată lumea vrut să ştie de ce autorităţile japoneze nu au putut preveni dezastrul.158 în toate aceste trei situaţii existau suficiente semnale de alarmă.

În ce priveşte inteligenţa artificială, există semnale clare de alarmă care ne arată că se pot produce crize în viitor, chiar dacă aceste semnale nu sunt neapărat evidente. Deşi există mai multe, iată două exemple care merită avute în vedere, precum şi potenţialele lor consecinţe.

Semnalul de alarmă nr. 1: Considerăm în mod eronat că inteligenţa artificială ar reprezenta o platformă digitală – asemănătoare cu internetul – care nu are principii de

dezvoltare sau planuri de creştere pe termen lung. Nu am recunoscut faptul că inteligenţa artificială a devenit o resursă publică. Definiţia economică a „bunurilor publice” e foarte strictă: ele trebuie să fie accesibile tuturor şi să poată fi folosite în acelaşi timp de persoane diferite. Serviciile publice, cum ar fi apărarea naţională, stingerea incendiilor sau colectarea gunoiului reprezintă bunuri publice. Aceste bunuri publice pot apărea şi în urma acţiunilor firmelor private, iar pe măsură ce timpul se scurge, bunurile publice produse de pieţe pot provoca consecinţe neprevăzute. În epoca noastră există un exemplu excelent al consecinţelor generalizării tehnologiei sub forma unei platforme: internetul.

Internetul a început sub forma unui concept – o modalitate de îmbunătăţirea a comunicării şi activităţii lucrative care urma în cele din urmă să aducă beneficii societăţii. Internetul modern a evoluat în urma unei colaborări de două decenii între diferiţi cercetători: la început era o reţea dezvoltată de departamentul Apărării, apoi a devenit o reţea mai vastă pe care cercetătorii o foloseau pentru a comunica între ei. Tim Berners-Lee, un programator de la CERN, a scris o propunere prin care reţeaua putea fi extinsă pe baza unui nou set de tehnologii şi protocoale care ar fi permis accesul altor persoane: localizatorul uniform de resurse (URL), limbajul hipertext (HTML) şi protocolul de transfer hipertext (HTTP). Reţeaua internetului a început să crească şi să aibă din ce în ce mai mulţi utilizatori pentru că era descentralizată, putea fi accesată de orice persoană care avea acces la un calculator iar utilizatorii noi nu-i împiedicau pe cei deja existenţi să creeze noi pagini.

Internetul nu a fost gândit iniţial ca un bun public şi nici nu a fost creat ca o reţea la care să aibă acces toţi locuitorii planetei, aşa cum este în prezent. De vreme ce nu a fost niciodată definit şi adoptat formal ca un bun public, el s-a supus în mod constant solicitărilor şi dorinţelor contrastante ale companiilor, agenţiilor

guvernamentale, universităţilor, militarilor, jurnaliştilor, directorilor din industria de film, militanţilor pentru drepturile omului dar şi oamenilor obişnuiţi din lumea întreagă. Ca urmare, s-au generat atât oportunităţi extraordinare cât şi rezultate imposibil de susţinut. În anul curent – 2019 – sărbătorim a cincizecea aniversare a momentului în care două calculatoare au făcut schimb de date pe o reţea mai largă, iar în umbra scandalului implicării ruseşti în alegerile prezidenţiale din Statele Unite şi a experimentelor psihologice efectuate de Facebook asupra a 700.000 de utilizatori fără cunoştinţa lor, unii dintre arhitecţii iniţiali ai internetului îşi doresc să fi luat decizii mai bune în urmă cu câteva decenii.159 Berners-Lee a lansat un apel, îndemnându-ne să rezolvăm toate problemele neprevăzute provocate de evoluţia internetului.160

Deşi numeroşi oameni inteligenţi susţin folosirea inteligenţei artificiale *pentru* binele public, nu vorbim încă despre inteligenţa artificială ca despre un bun public. Această abordare este eronată. Ne aflăm la începutul evoluţiei moderne a inteligenţei artificiale şi nu putem să continuăm să o considerăm o platformă construită de cei Nouă Titani pentru comerţ electronic, comunicare şi aplicaţii atrăgătoare. Dacă nu vom considera inteligenţa artificială un bun publicaşa cum considerăm de exemplu aerul pe care îl respirămvom întâlni probleme grave, de nerezolvat. Dacă considerăm că inteligenţa artificială reprezintă un bun public nu vom împiedica accesul Mafiei G la încasări sau dreptul membrilor lor de a creşte. Am produce doar o modificare a gândirii şi aşteptărilor, într-o bună zi, nu vom mai avea luxul de a dezbate şi de a discuta automatizarea în contextul drepturilor omului şi al geopoliticii. Inteligenţa artificială va fi devenit prea complexă pentru ca noi să o mai putem demonta sau modela după bunul nostru plac.

Semnalul de alarmă nr. 2: Inteligenţa artificială concentrează cu rapiditate puterea în jurul unor grupuri

restrânse, chiar dacă noi avem impresia că inteligenţa artificială reprezintă un ecosistem deschis cu puţine bariere. Viitorul inteligenţei artificiale este construit de către două ţări – Statele Unite şi Chinacare au interese geopolitice opuse, cu economii strâns legate una de cealaltă, şi cu lideri care exprimă adesea puncte de vedere divergente. Ca urmare, viitorul inteligenţei artificiale reprezintă un instrument prin care se manifestă atât puterea explicită cât şi cea subtilă, fiind manipulat – împreună cu triburile inteligenţei artificiale – pentru a obţine câştiguri economice şi avantaje strategice. Cadrul legislativ al celor două ţări ar putea fi corect, cel puţin pe hârtie, pentru viitorul maşinilor gânditoare. În lumea reală însă, acest cadru generează riscuri.

Filosofia capitalistă şi spiritul antreprenorial al Statelor Unite nu conduc întotdeauna la o creştere absolută şi la oportunităţi nelimitate. Ca şi în cazul oricăror alte industrii-telecomunicaţii, sănătate, fabricarea autoturismelor – cu timpul, competiţia se restrânge, puterea se consolidează în mâinile câtorva companii şi avem din ce în ce mai puţine variante dintre care putem alege pe măsură ce ecosistemul ajunge la maturitate. Putem alege dintre două sisteme de operare pentru telefoane mobile, iOS de la Apple, care are o cotă de piaţă de 44% în Statele Unite, şi Android de la Google, care are o cotă de 54%, în creştere. (Sub 1% dintre americani folosesc telefoane Microsoft sau Blackberry.) 161 Americanii au mai multe variante când vine vorba de serviciile de e-mail, dar 61% dintre cei cu vârste între 19 şi 34 de ani folosesc Gmail iar restul folosesc Yahoo şi Hotmail (19% respectiv 14%).162 Putem cumpăra de pe orice magazin online, dar cu toate acestea 50% dintre tranzacţiile electronice din Statele Unite sunt efectuate prin Amazon. Cei mai apropiaţi concurenţi – Walmart, Best Buy, Ma – deţin împreună o cotă de piaţă de sub 8%.163

Prin intermediul inteligenţei artificiale oricine poate construi un nou produs sau serviciu, dar nu îl poate

comercializa cu uşurinţă fără ajutorul Mafiei G. Ei trebuie să folosească tehnologia Tensor Flow a Google, diferiţii algoritmi de recunoaştere ai Amazon, serverele Azure de la Microsoft pentru găzduirea platformei, microchipurile IBM şi oricare alte structuri, instrumente şi servicii care fac sistemul să funcţioneze. În practică, viitorul inteligenţei artificiale nu se pregăteşte în Statele Unite într-o piaţă cu adevărat liberă.

Motivul pentru care această putere este atât de concentrată este că pentru ca inteligenţa artificială să ajungă unde se află acum a fost nevoie de o evoluţie de câteva decenii. Guvernul nostru ar fi trebuit să finanţeze mult mai mult inteligenţa artificială încă din anii ’80 şi ar fi trebuit să sprijine universităţile pentru a se putea pregăti pentru cea de-a treia epocă a informaticii. Spre deosebire de China, guvernul Statelor Unite nu a urmărit un plan integrat prin care să aloce sute de miliarde de dolari în acest domeniu-avansul a fost creat organic de sectorul privat. Aceasta înseamnă, implicit, că am cerut şi i-am permis Mafiei G să ia deciziile importante, semnificative care afectează viitorul forţei de muncă, siguranţei naţionale, creşterii economice şi al oportunităţilor noastre individuale.

Între timp, variantă chineză a comunismului – un socialism de piaţă combinat cu standarde clare în domeniul social – ar putea teoretic să stimuleze armonia şi stabilitatea politică, să sporească nivelul mediu de trai şi să împiedice un miliard de oameni să se răscoale. În practică, vorbim de un regim în întregime autoritar. Impactul asupra inteligenţei artificiale înseamnă un demers coordonat de a strânge cantităţi uriaşe de date despre cetăţeni, un sprijin acordat grupului BAT şi sporirea influenţei globale a Partidului Comunist Chinez.

E greu să ne putem imagina crizele şi oportunităţile înainte ca ele să apară, motiv pentru care avem tendinţa să preferăm explicaţiile pe care le cunoaştem. De aceea aducem în discuţie mai degrabă roboţi ucigaşi decât mici

tăieturi cu hârtie. De aceea fetişizăm viitorul inteligenţei artificiale în loc să ne temem de algoritmii care învaţă din datele noastre personale. Am descris doar două semnale de alarmă, dar există multe altele de care ar trebui să ţinem cont. Avem ocazia să recunoaştem atât beneficiile extraordinare dar şi riscurile potenţiale ale parcursului pe care se dezvoltă în prezent inteligenţa artificială. E şi mai important să ţinem cont de obligaţia de a acorda atenţie semnalelor de alarmă din prezent. Nu vrem să ne regăsim într-o situaţie în care va trebui să scuzăm sau să explicăm comportamentul inteligenţei artificiale aşa cum am făcut-o şi după scandaluri ca cele din Flint, naveta spaţială *Columbia* sau Fukushima.

Trebuie să căutăm în mod activ semnalele de alarmă şi să construim parcursuri alternative la traiectoria inteligenţei artificiale pentru a putea anticipa riscurile şi să sperăma evita o catastrofă, în momentul actual, nu există o metodă probabilistică prin care să putem prezice corect viitorul. Explicaţia se bazează pe faptul că oamenii sunt nişte fiinţe capricioase, că nu putem lua în calcul haosul şi hazardul, şi că numărul de date de care ar trebui să ţinem cont este în continuă creştere. Ca futurolog profesionist care foloseşte frecvent datele cantitative în cercetare, ştiu că deşi putem prevede rezultatul unui eveniment cu un set limitat de informaţii (cum ar fi alegerile), când vine vorba de inteligenţa artificială există un număr uriaş de variabile care trebuie detectate. Numărul de oameni care iau decizii în cadrul şedinţelor, când programează şi când aleg ce algoritmi să prelucreze anumite date e prea mare; există prea multe mici progrese care nu sunt publicate în jurnalele de specialitate; prea multe alianţe, achiziţii şi angajări pe care le fac cei Nouă Titani; prea multe proiecte de cercetare pe care le efectuează universităţile. Deşi nu putem să facem previziuni despre inteligenţa artificială, putem cu siguranţă să stabilim legături între semnalele de alarmă, semnalele slabe şi alte informaţii din prezent.

Am dezvoltat o metodologie de a modela incertitudinea profundă. Este un proces cu şase etape care scoate la iveală tendinţele emergente, identifică punctele comune şi legăturile dintre ele, le urmăreşte traiectoriile pe parcursul timpului, descrie potenţialele rezultate şi construieşte în cele din urmă o strategie pentru a obţine viitorul dorit. Prima jumătate a metodologiei descrie „ce-ul”, în vreme ce a doua jumătate descrie „cum ar fiul”. Această a doua jumătate se numeşte formal „planificarea scenariilor” dezvoltând scenarii despre viitor pe baza unei game vaste de date din diferite surse: statistici, cererile de patentare, cercetarea academică şi arhivistică, comunicate ale autorităţilor, articole prezentate la conferinţe, interviuri structurate cu diferite persoane şi chiar design critic şi ficţiune speculativă.

Planificarea scenariilor îşi are originile la începutul Războiului Rece, în anii ’50. Herman Kahn, un futurolog de la Corporaţia RÂND a primit sarcina să cerceteze războiul nuclear, iar el ştia că doar pe baza datelor nu va putea oferi suficiente informaţii liderilor armatei. El a dezvoltat un concept nou pe care l-a denumit „scenariu”. Aceste scenarii urmau să ofere detaliile descriptive şi povestea de care era nevoie pentru ca persoanele însărcinate cu elaborarea strategiei militare să poată înţelege rezultatele probabilerespectiv, ce s-ar putea întâmpla dacă ar fi luate un set anume de măsuri. În aceeaşi perioadă, futurologii francezi Bertrand de Jouvenel şi Gaston Berger au dezvoltat şi folosit scenarii prin care să descrie rezultatele preferate – ce *ar trebui* să se petreacă, date fiind împrejurările actuale. Studiile lor au forţat armata şi politicienii să ia în calcul „variantele de neconceput” şi consecinţele unui război nuclear. Exerciţiul de gândire a fost atât de reuşit încât şi alte guverne şi companii din lume au adoptat abordările lor. Compania Royal Dutch Shell a popularizat planificarea scenariilor atunci când a dezvăluit că scenariile i-au făcut pe directori să anticipeze criza globală a energiei (1973 şi 1979) precum şi

prăbuşirea bursei din 1986 şi să poată combate riscurile înaintea concurenţei.164 Scenariile reprezintă un instrument atât de puternic încât Shell are şi după 45 de ani o echipă mare dedicată cercetării şi elaborării acestor scenarii.

Am pregătit scenarii de riscuri şi oportunităţi legate de viitorul inteligenţei artificiale pentru multe domenii şi industrii şi un grup divers de organizaţii. Scenariile sunt un instrument care ne ajută să combatem o prejudecată cognitivă pe care savantul economist şi de drept Cass Sunstein o denumeşte „neglijarea probabilităţii”. 165 Creierele umane nu se pricep să evalueze riscurile şi pericolele. Presupunem că activităţile considerate normale sunt mai sigure decât cele noi sau mai puţin frecvente. Spre exemplu, majoritatea consideră că a conduce o maşină e mult mai sigur decât a zbura cu un avion de linie, deşi călătoria cu avionul reprezintă cel mai sigur mod de transport. Americanii au o şansă de 1 la 114 de a muri într-un accident de maşină, însă o şansă de 1 la 9821 de a muri într-un avion.166, 167 Evaluăm incorect riscul de a conduce o maşină, motiv pentru care mulţi oameni scriu mesaje pe telefon în timp ce conduc sau consumă băuturi alcoolice. Suntem la fel de nepricepuţi în a evalua riscurile pe care le presupune inteligenţa artificială pentru că o folosim automat în fiecare zi fără să ne dăm seama, în timp ce apreciem şi partajăm postări, trimitem e-mailuri şi mesaje text, vorbim cu roboţi şi ne lăsăm influenţaţi. Orice risc pe care ni-l imaginăm provine din science fiction: inteligenţa artificială sub forma unor androizi care vânează oameni sau voci distorsionate care ne torturează psihologice. Nu ne gândim în mod natural la viitorul inteligenţei artificiale în cadrul capitalismului, geopoliticii şi democraţiei. Nu ne imaginăm şinele din viitor şi felul în care aceste sisteme autonome ne-ar putea afecta sănătatea, relaţiile şi fericirea.

Am avea nevoie de un set de scenarii orientate către public care să descrie toate modurile prin care inteligenţa

artificială şi cei Nouă Titani ne pot afecta colectiv, pe măsură ce inteligenţa artificială trece de la aplicaţii ţintite la sisteme dotate cu inteligenţă generală şi chiar mai mult decât atât. Am trecut de punctul de inacţiune. Gândiţi-vă astfel: în apa de băut e plumb. Garniturile sunt putrede. În reactor există crăpături. Starea actuală a inteligenţei artificiale a generat probleme sistemice fundamentale pentru care există semnale de alarmă, iar aceste probleme trebuie rezolvate acum. Dacă luăm astăzi măsurile corecte care se impun, ne aşteaptă nişte oportunităţi extraordinare în viitor.

În următoarele capitole voi prezenta în amănunt trei scenarii – unul optimist, unul pragmatic şi unul catastrofal – pe care le-am modelat folosind date şi amănunte din zilele noastre. Ele intră în domeniul speculativ dar sunt bazate pe fapte concrete. Scopul acestor scenarii este de a da un caracter de iminenţă şi realitate unor noţiuni care par îndepărtate sau fantasmatice. Pentru că nu putem vedea cu uşurinţă cum acţionează inteligenţa artificială, remarcăm rezultatele doar atunci când sunt negativeiar în acel moment oamenilor obişnuiţi nu le rămân prea multe căi de atac.

Drumul de la inteligenţa artificială îngustă la super-inteligenţa artificială

Prima parte a cărţii a analizat în principal inteligenţa artificială îngustă şi automatizarea a milioane de aplicaţii zilnice – de la depistarea cecurilor frauduloase la evaluarea candidaţilor pentru slujbe sau stabilirea preţului biletelor de avion. Pentru a parafraza cuvintele renumitului informatician Frederick Brooks de la IBM, nu puteţi construi programe din ce în ce mai complexe doar sporind numărul oamenilor care lucrează la rezolvarea problemelor. Sporirea numărului de programatori are tendinţa de a încetini evoluţia proiectelor.168 în prezent, oamenii trebuie să proiecteze sisteme şi să scrie programe pentru a evolua diferite aplicaţii ale inteligenţei artificiale, existând la fel ca în orice alt domeniu de cercetare o curbă

de învăţare. Acesta este motivul pentru care atingerea următoarei etape de evoluţie a inteligenţei artificiale este atât de atrăgătoare pentru cei Nouă Titani. Sistemele care se pot programa singure ar putea procesa mult mai multe date, ar putea construi şi testa noi modele şi s-ar putea îmbunătăţi singure fără nevoia unei implicări directe a oamenilor.

Inteligenţa artificială se încadrează de regulă într-una din următoarele trei categorii: inteligenţa artificială îngustă sau slabă (L-AI), inteligenţa artificială generală (IAG) şi super-inteligenţa artificială (SLA). Cei Nouă Titani lucrează în prezent la construirea şi exploatarea unor sisteme IAG pe care speră să le poată folosi la un moment dat pentru a raţiona, a rezolva probleme, a gândi abstract şi a face alegeri la fel de uşor ca noi, cu rezultate comparabile sau mai bune. IAG aplicată ar însemna obţinerea unor rezultate mult mai rapide în cercetare, dar şi îmbunătăţirea diagnosticelor medicale sau găsirea unor noi metode prin care să se rezolve probleme dificile de inginerie. îmbunătăţirile LAG ar trebui în cele din urmă să conducă la apariţia ultimei categorii-super-inteligenţa artificială. Sistemele SLA pot varia de la a avea o performanţă uşor mai bună la nivel cognitiv decât oamenii până la sisteme care să fie de mii de miliarde de ori mai inteligente decât oamenii, sub orice aspect.

Tranziţia de la momentul în care ne aflăm în zilele noastre la folosirea pe scară largă a IAG presupune folosirea unor „algoritmi evolutivi”, un domeniu de cercetare inspirat de scrierile lui Charles Darwin privind selecţia naturală. Darwin a descoperit că cei mai puternici membri ai unei specii supravieţuiesc cu timpul, transmiţând cu precădere moştenirea lor genetică în cadrul populaţiei. Cu timpul, specia devine mai bine adaptată la mediul său. Şi inteligenţa artificială funcţionează la fel. Un sistem porneşte cu un set semialeatoriu sau aleatoriu de posibilităţi (vorbim de miliarde sau de mii de miliarde de informaţii) şi rulează

simulări. Deoarece soluţiile generate iniţial sunt aleatorii, ele nu pot fi aplicate cu succes în lumea reală, însă există şansa ca unele să fie un pic mai bune decât altele. Sistemul va elimina variantele mai slabe şi le va reţine pe cele mai puternice, creând noi combinaţii. Uneori, combinaţiile noi vor genera soluţii hibride, care vor fi de asemenea păstrate. Tot uneori, o schimbare aleatorie va provoca o mutaţie – similară evoluţiei organice a oricărei specii. Algoritmul evolutiv va continua să genereze, elimine şi promoveze soluţii de milioane de ori producând mii şi chiar milioane de copii, până când va ajunge într-un final la concluzia că nu mai poate îmbunătăţi nimic. Algoritmii evolutivi cu puterea de a se schimba vor ajuta inteligenţa artificială să evolueze singură iar această posibilitate este una ofertantă, dar care vine la pachet cu un cost: felul în care funcţionează soluţia rezultată şi procesul prin care s-ar ajunge la acea soluţie ar putea fi prea complex pentru ca să poată fi înţeles şi interpretat chiar şi de către cei mai inteligenţi programatori.

Acesta este motivul pentru care e important – chiar dacă pare fantezist – ca maşinile să fie incluse în orice discuţie legată de evoluţia speciei noastre. Până în prezent, ne-am gândit la evoluţia vieţii pe Pământ folosind un interval limitat. Acum sute de milioane de ani, organismele uni-celulare au înghiţit alte organisme şi au dat naştere unor noi forme de viaţă. Procesul a continuat până când primii oameni au putut să stea în două picioare, au evoluat genunchi, s-au adaptat la mersul biped, au căpătat femururi mai lungi şi şi-au dat seama cum să construiască topoare şi să controleze focul, au evoluat creiere mai mari şi în cele din urmă după milioane de selecţii darwiniene naturaleau putut construi primele maşini gânditoare. Ca şi roboţii, şi corpurile noastre sunt doar nişte receptacule ale unor algoritmi complecşi. Trebuie să ne gândim la evoluţia vieţii ca la o evoluţie a inteligenţei: inteligenţa omenească şi inteligenţa artificială au urmat parcursuri paralele într-un ritm care ne-a păstrat

locul întâi în ierarhia inteligenţei. Şi asta în ciuda criticilor conform cărora generaţiile viitoare vor deveni mai proaste din cauza tehnologiei. Îmi aduc aminte că profesorul de matematică din liceu era nemulţumit de calculatoarele de buzunar care fuseseră lansate pe piaţă cu doar cinci ani înainte iar el susţinea că generaţia noastră devenea mai proastă şi mai leneşă. Deşi susţinem că generaţiile viitoare vor fi mai proaste *din cauza* tehnologiei, nu ne gândim niciodată că oamenii s-ar putea trezi într-o bună zi mai proşti *decât* tehnologia. Ne apropiam de această cotitură, iar ea este legată de limitele evoluţiei noastre.

Inteligenţa omenească este măsurată cel mai adesea prin intermediul unei metode dezvoltate în 1912 de psihologul german William Stern. Cunoaşteţi această metodă sub numele de „coeficient de inteligenţă” sau IQ. Scorul se calculează împărţind rezultatul unui test de inteligenţă la vârsta exprimată în ani apoi înmulţind rezultatul cu 100. Aproximativ 2, 5% din populaţie obţine un scor de peste 130, fiind consideraţi gânditori de elită, în vreme ce 2, 5% au un scor sub 70, fiind consideraţi a avea dizabilităţi de învăţare sau psihice. Chiar şi cu câteva procentaje ca marjă de eroare, două treimi din populaţie se încadrează între 85 şi 115 pe această scară. Cu toate acestea, suntem ceva mai inteligenţi decât eram în trecut. De la începutul secolului XX, scorul mediu creşte într-un ritm de aproximativ 3 puncte în fiecare deceniu, probabil datorită îmbunătăţirii alimentaţiei, educaţiei şi complexităţii mediului.169 Ca urmare, nivelul mediu al oamenilor a început să crească. Dacă această tendinţă va continua, ne-am aştepta să avem mai multe genii până la finalul secolului decât avem acum. Între timp, evoluţia noastră biologică se va intersecta cu inteligenţa artificială.

Pe măsură ce capacitatea noastră intelectuală se îmbunătăţeşte, se va îmbunătăţi şi cea a inteligenţei artificiale – însă nu putem folosi testele IQ pentru a măsura inteligenţa artificială. Folosim în schimb numărul de operaţii sau calcule pe secundă, cunoscute şi sub

numele de ops, pe care le poate efectua un calculator, pe care le putem compara apoi cu creierul uman. În funcţie de sursă, numărul maxim de operaţii pe secundă pe care le pot efectua creierele noastre este de 1 exaflop, aproximativ un miliard de miliarde de operaţiuni pe secundă, iar acestea reflectă multe activităţi care se petrec fără să fim conştienţi de ele: micile mişcări pe care le facem când respirăm, procesarea vizuală continuă care se petrece când avem ochii deschişi şi aşa mai departe. Lansat în 2010, computerul chinezesc Tianhe-l era cel mai rapid şi mai puternic supercalculator, construit în întregime cu microprocesoare chinezeşti şi care poate atinge un maxim teoretic de 1, 2 petaflops. (Un petaflop înseamnă un milion de miliarde de operaţii pe secundă.) E rapid, dar nu se apropie de creierul omenesc. Apoi, în luna iunie 2018, IBM şi Departamentul Energiei au lansat computerul Summit, care poate atinge 200 de petaflops, fiind construit în mod specific pentru inteligenţa artificială. 170 Această înseamnă că ne apropiem de o maşină cu o putere de calcul mai mare decât a noastră, chiar dacă ea nu ar putea încă trece testul Turing şi să ne păcălească să credem că e un om.

Viteza nu este însă unicul parametru relevant. Dacă am accelera viteza creierului unui câine la I015 ops, el nu ar putea dintr-odată să se apuce să rezolve ecuaţii diferenţialeel doar ar alerga prin curte şi ar mirosi şi urmări mai multe obiecte. Creierul omului are o arhitectură mai complexă decât a unui câine: avem mai multe conexiuni între neuroni, proteine speciale şi noduri cognitive complicate. 171 Chiar şi aşa, inteligenţa artificială poate fi extinsă de o manieră imposibilă pentru oameni fără a schimba arhitectura de bază a creierului nostru. Legea lui Moore, care susţine că numărul de componente din circuitele integrate se dublează la fiecare doi ani pe măsură ce tranzistorii se micşorează continuă să se aplice şi ne spune că evoluţia computerelor urmează un tipar de creştere exponenţială. Există din ce în ce mai

multe date, precum şi noi algoritmi, componente avansate sau noi modalităţi de a conecta reţelele neurale. Toate aceste evoluţii generează mai multă putere. Spre deosebire de computere, noi nu putem modifica uşor structura creierelor noastre şi arhitectura inteligenţei omeneşti. Pentru a putea face asta ar trebui să (1) înţelegem în totalitate cum funcţionează creierele noastre (2) să modificăm arhitectura şi substanţele chimice din creierele noastre iar aceste modificări să poată fi transmise generaţiile următoare şi (3) să aşteptăm o perioadă lungă de timp până când vom face copii.

În ritmul actual, pentru ca IQ-ul oamenilor să crească cu 15 puncte e nevoie de 50 de ani. Iar pentru noi 15 puncte se observă imediat. Diferenţa dintre un creier din partea superioară a mediei, cu un IQ de 119, şi unul cu un IQ de 134 „dotat” ar presupune o capacitate cognitivă semnificativ mai mareefectuarea cu mai mare rapiditate a conexiunilor, stăpânirea mai uşoară a noilor concepte şi o gândire mai eficientă. În acelaşi interval de timp, capacităţile cognitive ale inteligenţei artificiale nu doar că ne vor depăşiele vor deveni de nerecunoscut pentru noi pentru că nu avem din punct de vedere biologic suficientă putere de calcul pentru a înţelege ce este inteligenţa artificială. Pentru noi, o întâlnire cu o maşină super-inteligentă ar fi ca un cimpanzeu care ar participa la o şedinţă de consiliu local. Cimpanzeul ar putea să recunoască faptul că în cameră sunt oameni şi că poate sta pe scaun, dar ce ar înţelege dintr-o discuţie despre benzi de bicicletă într-o intersecţie aglomerată? El nu dispune nici pe de parte de capacitatea cognitivă de a descifra limbajul folosit, darmite de argumentele şi experienţa care i-ar permite să combată propunerea benzilor de bicicletă. Pe termen lung în evoluţia inteligenţei şi drumul către să-i, noi oamenii jucăm rolul cimpanzeului.

O inteligenţă artificială super-inteligentă nu e neapărat periculoasă, şi nu subminează neapărat rolul pe care îl jucăm în cadrul civilizaţiei. Cu toate aceste, SIA ar

lua decizii de o manieră inconştientă folosind o logică de care am fi străini. Filosoful Nick Bostrom de la Universitatea Oxford explică rezultatele posibile ale SIA folosind o parabolă cu agrafe de hârtie. Dacă am cere unei SĂI să fabrice agrafe, ce s-ar petrece? Rezultatele fiecărei inteligenţe artificiale, inclusiv cele care există deja, sunt stabilite prin valori şi obiective. E posibil ca SIA să inventeze o nouă agrafa, mai bună, care să ţină hârtiile chiar şi când ar cădea pe jos. E posibil că dacă nu suntem capabili să spunem câte agrafe ne *trebuie*, SIA să continue să fabrice la infinit agrafe, umplându-ne casele şi birourile cu ele, dar şi spitalele, şcolile, râurile, lacurile, canalizarea şi aşa mai departe până când întreaga planetă e plină de agrafe. O SIA care acordă prioritate eficienţei ar putea decide că oamenii stau în calea agrafelor, modificând întreaga planetă şi transformând-o într-o fabrică de agrafe, provocând dispariţia speciei noastre. 172 Iată de ce mulţi experţi în inteligenţă artificială, printre care mă număr şi eu, sunt îngrijoraţi: dacă abilităţile cognitive ale SIA sunt de multe ori mai bune decât ale noastre (reţineţi, nu suntem decât ceva mai inteligenţi decât cimpanzeii) atunci suntem incapabili să ne imaginăm consecinţele pe care le-ar putea avea aceste maşini inteligente asupra civilizaţiei noastre.

Acesta este motivul pentru care cercetătorii din domeniul inteligenţei artificiale folosesc frecvent cuvântul „explozie”. Matematicianul şi criptologul britanic I.J. Good a folosit prima dată termenul într-un eseu din 1965: „O maşină ultrainteligentă ar putea crea maşini şi mai performante; s-ar produce indenegabil o „explozie a inteligenţei” iar inteligenţa omenească ar rămâne cu mult în urmă. Astfel, prima maşină ultrainteligentă ar fi ultima invenţie pe care omul ar trebui să o mai facă, atât timp cât maşina e suficient de docilă încât să ne spună cum să o ţinem sub control.173

Cei Nouă Titani construiesc structuri şi sisteme care – ei sperăvor încuraja într-o bună zi o astfel de explozie,

creând un teren propice unor soluţii, strategii, concepte, structuri şi abordări complet noi, la care nici măcar cei mai inteligenţi dintre informaticieni nu s-au gândit. Astfel ar urma evoluţii din ce în ce mai rapide, oportunităţi şi creşterea cifrei de afaceri. În termeni de specialitate, o astfel de explozie se numeşte „auto-îmbunătăţire recursivă”, referindu-se la un ciclu prin care inteligenţa artificială devine prin propriile forţe mai rapidă, mai bună şi mai inteligentă prin modificarea capacităţilor sale. Astfel, inteligenţa artificială ar putea să preia controlul şi să-şi făurească propriul destin, îmbunătăţirile s-ar putea produce la fiecare oră, sau chiar instantaneu.

„Explozia inteligenţei” care va urma descrie nu doar viteza super-computerelor sau puterea algoritmilor ci şi răspândirea la scară largă a maşinilor gânditoare inteligente care urmăresc autoîmbunătăţirea recursivă. Imaginaţi-vă o lume în care sisteme mult mai avansate decât Alpha Go Zero şi NASNet nu doar că iau decizii strategice în mod autonom, ci şi colaborează şi concurează ca parte a unei comunităţi globale. O lume în care li se cere să evolueze, în principal pentru a ajuta oameniiscriind noi generaţii de programe, evoluând şi auto-îmbunătăţindu-se, însă la o viteză uluitoare. Inteligenţele artificiale ar crea noi agenţi, programându-i cu un scop şi un set de sarcini, iar ciclul s-ar repeta în continuu, de mii de miliarde de ori, provocând atât schimbări minore cât şi colosale. Singura epocă istorică în care s-a produs un cataclism evolutiv de o magnitudine similară a avut loc aproximativ acum 542 de milioane de ani în epoca cambriană, când diversificarea rapidă a formelor de viaţă a condus la apariţia unor noi fiinţe complexe şi a transformat pentru totdeauna planeta noastră. Fostul director de programe de la DARPA Gill Pratt crede că ne aflăm în mijlocul unei explozii de tip cambriano perioadă în care inteligenţa artificială învaţă din experienţa tuturor inteligenţelor artificiale, după care viaţa pe Pământ ar putea fi complet diferită faţă de prezent.174

Acesta e motivul pentru care cei Nouă Titani, investitorii şi acţionarii lor, agenţiile guvernamentale şi politicienii, cercetătorii trebuie să recunoască semnalele de alarmă şi să se gândească mai critic nu doar la viitorul IAÎ care este creată acum, dar şi la LAG şi SIA care se întrevăd la orizont. Şi voi jucaţi un rol la fel de important în cadrul acestui proces. Evoluţia inteligenţei constituie un spectru, pe care coexistă atât oamenii cât şi maşinile. Valorile celor Nouă Titani sunt deja profund întipărite în algoritmii, sistemele şi structurile existente. Aceste valori vor fi transmise mai departe către milioane de noi generaţii de inteligenţe artificiale care evoluează, şi în curând unor maşini gânditoare capabile de inteligenţă generală.

Tranziţia de la IAÎ la SIA va dura probabil următorii 70 de ani. În momentul de faţă sunt dificil de definit etapele concrete pentru că progresul în domeniul inteligenţei artificiale depinde de mai mulţi factori şi oameni: noii membri incluşi în triburile inteligenţei artificiale, deciziile strategice luate de cei Nouă Titani, disputele comerciale şi geopolitice, ca să nu mai vorbim de evenimentele haotice sau de hazard. În propriile mele modele, consider că IAG va apărea în jurul anului 2040. Pare că e un viitor îndepărtat, aşa că voi încerca să creez un context pentru această dată. Vom fi avut trei sau patru preşedinţi diferiţi la Casa Albă. (în absenţa unor probleme de sănătate, preşedintele Xi Jinping va fi în continuare la putere.) Voi avea 65 de ani când sistemele IAG vor putea efectua propria cercetare în domeniul inteligenţei artificiale. Copilul meu de clasa a doua va avea 30 de ani, iar în acel moment ar putea să citească un bestseller New York Times scris în întregime de o maşină. Tata va avea peste 90 de ani iar toţi specialiştii (cardiologi, nefrologi, radiologi) vor fi IAG, dirijate şi controlate de un medic generalist hiperspecializat, care va fi atât un medic cât şi un analist de date. Apariţia SIA ar putea urma la scurt sau mult timp după aceea, undeva între 2040 şi 2060. Nu e

obligatoriu ca până în 2070 SIA să fi distrus viaţa pe Pământ sub miliarde de agrafe. Cu toate acestea, nici această variantă nu poate fi exclusă.

Poveştile pe care trebuie să ni le spunem

Pentru a putea planifica viitorul inteligenţei artificiale avem nevoie de noi poveşti folosind date din lumea reală. Dacă suntem de acord că inteligenţa artificială va evolua pe parcurs, va trebui să construim scenarii care să descrie intersectarea celor Nouă Titani cu forţele economice şi politice care îi ghidează, precum şi rolul omenirii pe măsură ce inteligenţa artificială trece de la etapa aplicaţiilor înguste la inteligenţă generală şi în cele din urmă la super-inteligenţă.

Pentru că viitorul nu s-a produs încă, nu putem şti cu siguranţă toate rezultatele posibile ale acţiunilor noastre din prezent. Din acest motiv, scenariile următoare sunt scrise folosind diferite încadrări afective pentru următorii cincizeci de ani. Primul este scenariul optimist, în care ne întrebăm ce se întâmplă dacă cei Nouă Titani decid să adopte schimbări radicale pentru a se asigura că inteligenţa artificială va fi în beneficiul tuturor. Am o precizare importantă de făcut: scenariile „optimiste” nu sunt neapărat încurajatoare sau motivante. Ele nu conduc întotdeauna la o utopie. Într-un scenariu optimist presupunem că se iau cele mai bune decizii şi că se îndepărtează obstacolele care stau în calea reuşitei, în ce ne priveşte, aceasta ar presupune ca cei Nouă Titani să-şi schimbe abordarea privind inteligenţa artificială, iar pentru că ei iau tot timpul cele mai bune decizii, viitorul nostru va fi mult mai bun decât prezentul. E un scenariu în care mi-ar plăcea să trăiesc, şi un viitor pe care-l putem avea dacă lucrăm împreună.

Următorul este un scenariu pragmatic care descrie cum ar arăta viitorul dacă cei Nouă Titani nu fac pe termen scurt decât mici îmbunătăţiri. Presupunem că deşi toate persoanele interesate recunosc că inteligenţa artificială nu se îndreaptă într-o direcţie bună, nu va exista

o colaborare pentru a pune bazele unor schimbări semnificative şi de durată. Câteva universităţi vor introduce cursuri obligatorii de etică, Mafia G formează parteneriate la nivel de industrie pentru a combate riscurile dar fără să-şi modifice cultura organizaţională a membrilor; politicienii se concentrează asupra alegerilor şi pierd din vedere planurile ambiţioase ale Chinei. Scenariul pragmatic nu este unul al marilor schimbări-el ţine cont de fluxul şi refluxul demersurilor de îmbunătăţire ale omenirii. Acest scenariu ţine cont de faptul că liderii din mediul politic şi de afaceri sunt cât se poate de dispuşi să saboteze viitorul de dragul unor câştiguri imediate, pe termen scurt.

În ultimul rând, scenariul catastrofal explică ce s-ar întâmpla dacă toate semnalele de alarmă ar fi ignorate şi nu ne-am construi un plan pentru viitor, iar cei Nouă Titani vor continua să concureze unii împotriva celorlalţi. Dacă alegem să mergem mai departe pe aceleaşi baze, încotro am putea ajunge? Ce s-ar întâmpla dacă inteligenţa artificială şi-ar continua parcursul actual în Statele Unite şi China? Crearea unei schimbări sistemice – care ar putea împiedica scenariul catastrofale o muncă grea, de durată, şi fără un final concret. Acesta este motivul pentru care scenariul catastrofal pare atât de înfricoşător şi amănuntele atât de tulburătoare, fiindcă în prezent pare că ne îndreptăm către acest scenariu.

Am studiat, modelat şi conceput aceste situaţii pentru a descrie rezultatele începând cu anul 2029. La baza scenariilor se regăsesc câteva teme principale, printre care se numără: mobilitatea şi oportunităţile economice, productivitatea forţei de muncă, îmbunătăţirea structurilor sociale, dinamica de putere a celor Nouă Titani, relaţiile dintre Statele Unite şi China şi expansiunea sau retragerea globală a comunismului şi democraţiei. Voi arăta cum valorile noastre sociale şi culturale s-ar putea schimba pe măsură ce inteligenţa artificială ajunge la maturitate: felul în care definim creativitatea, felul în care

relaţionăm unii cu ceilalţi şi ce credem despre viaţă şi moarte. Deoarece scenariile urmăresc să ne ajute să înţelegem cum ar arăta viaţa noastră în timpul tranziţiei de la IAÎ la SLA, am inclus exemple legate de casele, locurile de muncă, educaţie, sănătate, ordine publică, urbanism, infrastructură locală, politică şi siguranţă naţională.

Unul dintre rezultatele probabile ale inteligenţei artificiale şi un element comun al tuturor celor trei scenarii îl reprezintă apariţia unui „Dosar de Date Personale”, sau DDP. Acesta constituie un registru unic care ar cuprinde toate datele generate de activitatea noastră digitală (gândiţi-vă la internet sau la telefoane mobile) dar ar include şi date din alte surse: notele de la şcoală, istoricul de muncă (diplome, foşti şi actuali angajatori); date cu caracter juridic (căsătorii, divorţuri, arestări); date financiare (ipoteci, scoruri de credit, împrumuturi, impozite); date despre călătorii (ţări vizitate, vize); istoricul relaţiilor (aplicaţii online); date de sănătate (arhive electronice de sănătate, analize genetice, activitatea sportivă) şi istoricul de cumpărături (magazine online, folosirea cupoanelor de reduceri). În China un astfel de DDP ar include şi datele privind scorul de credit social amintit în capitolul anterior. Inteligenţa artificială creată de către cei Nouă Titani ar putea atât să înveţe despre voi pe baza acestui DDP cât şi să îl folosească pentru a lua decizii automate şi a vă oferi o gamă de servicii. DDP ar putea fi moştenibil – un istoric cuprinzător transmis copiilor pentru a-l putea folosişi ar putea fi gestionat temporar sau proprietatea exclusivă a unei companii din cadrul celor Nouă Titani. DDP joacă un rol important în scenariile pe care urmează să le citiţi.

Aceste DDP nu există încă, însă din punctul meu de vedere există deja semnale care par să arate că ne îndreptăm către un viitor în care multiplele surse de date personale vor fi adunate laolaltă într-un singur fişier pus la dispoziţie şi actualizat de către cei Nouă Titani. De fapt, luaţi deja parte la acest sistem şi folosiţi o variantă

prototip de DDP. E vorba de adresa voastră de e-mail.

Adresa de e-mail a unei persoane obişnuite a fost făcută să acţioneze şi ca o modalitate de autentificare; numărul de telefon este folosit pentru validarea tranzacţiilor iar telefonul este folosit pentru a-i găsi în lumea fizică. Dacă folosiţi Gmail, Google – şi prin extensie, şi sistemele lor de inteligenţă artificială, această companie ştie mai multe despre voi decât ştiu partenerii voştri. Sistemul cunoaşte numele şi adresele de e-mail ale tuturor persoanelor cu care corespondaţi, precum şi informaţii demografice legate de ele (vârstă, gen, localitate de domiciliu). Google ştie când aveţi tendinţa să vă verificaţi e-mailul şi în ce împrejurări. Pe baza e-mailului ştie unde călătoriţi, de ce sume de bani dispuneţi şi ce cumpăraţi. Dacă faceţi poze cu un telefon Android, acesta recunoaşte feţele prietenilor şi rudelor, putând detecta abaterile de la tipar pentru a trage concluzii: de exemplu multe poze noi cu aceeaşi persoană ar putea însemna că aveţi o nouă prietenă (sau o relaţie extraconjugală). Google ştie de toate întâlnirile, programările la doctor şi planurile voastre de a vă duce la sală. Google ştie dacă ţineţi Ramadanul, Rosh Hashanah, dacă vă duceţi la biserică sau dacă nu vă interesează nicio religie. Google ştie unde ar trebui să fiţi de regulă marţi după-amiază, chiar dacă sunteţi în altă parte. Sistemul ştie ce căutaţi, folosind degetele sau vocea, aşa că-şi poate da seama dacă aveţi prima dată un avort spontan, dacă învăţaţi cum să gătiţi paella, dacă aveţi probleme de identitate sexuală, dacă vă trece prin cap să renunţaţi la carne sau să vă căutaţi o nouă slujbă. Aceste date sunt interconectate, iar sistemul învaţă de pe urma lor, monetizând şi transformând în produse concluziile în timp ce vă sugestionează către anumite direcţii prestabilite.

În prezent, Google are acces la aceste informaţii în mod voluntar doar pe baza unui singur produsădresa de Gmailpe care probabil o folosiţi şi pentru a cumpăra produse de pe Amazon sau a vă autentifica în Facebook.

Nu mă plâng, ci doar descriu viaţa modernă. Pe măsură ce inteligenţa artificială progresează, un DDP mai explicit va permite celor Nouă Titani o mai mare eficienţă, astfel încât ne vor sugestiona să acceptăm şi să adoptăm aceste DDP, chiar dacă nu înţelegem pe deplin ce presupune folosirea lor. Desigur că în China aceste DDP au fost deja introduse ca proiecte pilot în cadrul sistemului de credite sociale.

„Ne spunem poveşti pentru a putea trăi”, scria Joan Didion în cartea sa *The White Album*. „Interpretăm ceea ce vedem, şi alegem cele mai fezabile dintre variante”. Trebuie să avem de ales când vine vorba de inteligenţa artificială. A venit vremea să folosim informaţia disponibilă pentru a ne spune poveşti-scenarii care să descrie cum am putea convieţui cu maşinile gânditoare.

CAPITOLUL CINCI
PROSPERITATE ÎN A TREIA EPOCĂ A INFORMATICII: SCENARIUL OPTIMIST

Suntem în anul 2023 şi am luat cele mai bune decizii în privinţa inteligenţei artificialeam modificat parcursul de dezvoltare al inteligenţei artificiale, construim viitorul şi vedem deja schimbări permanente în bine. Triburile inteligenţei artificiale, universităţile, cei Nouă Titani, agenţiile guvernamentale, investitorii, cercetătorii şi oamenii obişnuiţi au ţinut cont de semnalele timpurii de alarmă.

Înţelegem că există o singură schimbare care va rezolva problemele deja create şi acum cea mai bună strategie este de a ne ajusta aşteptările legate de viitorul inteligenţei artificiale. Recunoaştem că inteligenţa artificială nu este doar un produs al Silicon Valleyun bun care să fie valorificat cât timp piaţa este în creştere.

Înainte de orice, recunoaştem de ce China a investit strategic în inteligenţa artificială şi cum parcursul de dezvoltare al inteligenţei artificiale se potriveşte cu rolul pe care China şi-l asumă în lumea viitorului. China nu încearcă să modifice balanţa comercială ci să obţină un

avantaj absolut asupra Statelor Unite din punct de vedere al puterii economice, dezvoltării forţei de muncă, influenţei geopolitice, forţei militare, influenţei sociale şi al politicilor de mediu. Dându-şi seama de acest lucru, politicienii noştri au construit, cu sprijinul deplin al Mafiei G şi al triburilor inteligenţei artificiale, o coaliţie internaţională care să protejeze şi să păstreze inteligenţa artificială ca un bun public. Această coaliţie presează China şi foloseşte pârghiile economice pentru a contracara folosirea inteligenţei artificiale cu scopul de a supraveghea cetăţenii şi a consolida puterea regimului comunist.

Recunoscând că China foloseşte inteligenţa artificială pentru a-şi atinge obiectivele economice şi militare în timp ce răspândeşte idealurile comunismului şi-şi consolidează puterea asupra societăţii, guvernul federal al Statelor Unite alocă fonduri substanţiale pentru a sprijini dezvoltarea inteligenţei artificiale, ceea ce reduce presiunea asupra Mafiei G de a căuta profituri rapide. Folosind cursa spaţială a anilor ’50 drept precedent, devine evident cât de uşor poate fi depăşită America de alte ţări în absenţa unei coordonări la nivel naţional. Devine de asemenea cât se poate de evident ce fel de influenţă poate să exercite America în domeniul ştiinţei şi tehnologiei atunci când avem o strategie naţională coordonatăgraţie guvernului federal au apărut tehnologii precum GPS sau internetul.

Nici inteligenţa artificială şi nici finanţarea nu sunt politizate şi toată lumea este de acord că reglementarea Mafiei G sau inteligenţei artificiale reprezintă o abordare greşită. O reglementare strictă, dură ar deveni desuetă chiar din momentul în care ar fi adoptată, ar sufoca inovaţia şi ar fi dificil de pus în practică. Cu sprijinul ambelor partide politice, americanii sprijină sporirea cheltuielilor federale în domeniul inteligenţei artificiale folosind strategia chineză drept model. Se alocă fonduri pentru cercetare şi dezvoltare, studii de impact asupra economiei şi forţei de muncă, studii de impact social,

programe destinate diversităţii, iniţiative medicale şi de sănătate publică şi infrastructură şi pentru revitalizarea sistemului public de educaţie, oferind salarii atractive profesorilor şi o programă care oferă tuturor o mai bună pregătire pentru un viitor din ce în ce mai automatizat. Nu mai presupunem că Mafia G poate să-şi servească în egală măsură stăpânii politici şi financiari, sau că pieţele libere şi spiritul nostru antreprenorial vor produce cele mai bune rezultate pentru inteligenţa artificială şi pentru omenire.

Având la dispoziţie o strategie naţională şi finanţare, nou înfiinţata coaliţie a Mafiei G semnează acorduri oficiale multilaterale pentru a colabora la dezvoltarea viitorului inteligenţei artificiale. Coaliţia Mafiei G defineşte şi adoptă standarde care acordă prioritate dincolo de orice unui parcurs de dezvoltare pentru inteligenţa artificială care ţine cont de interesele democraţiei şi societăţii. Coaliţia acceptă unirea tehnologiilor din domeniul inteligenţei artificiale. Prin colaborare rezultă chipuri, structuri şi arhitecturi de reţea superioare, care înlocuiesc sistemele rivale de inteligenţă artificială şi o comunitate de programatori cu mai multe obiective. Această colaborare permite şi cercetătorilor să caute noi oportunităţi, iar astfel toată lumea câştigă.

Coaliţia Mafiei G adoptă transparenţa ca valoare fundamentală, rescriind în mod radical reglementările, condiţiile şi fluxurile de lucru pentru a acorda prioritate înţelegerii şi educaţiei. Acest proces se produce voluntar, evitând astfel intervenţia autorităţilor. Seturile de date, algoritmii de pregătire şi structurile de reţele neurale sunt transparentizate de o manieră care să protejeze doar acele secrete comerciale şi informaţii confidenţiale care ar putea provoca pagube economice unuia dintre membrii coaliţiei. Departamentele juridice ale Mafiei G nu mai petrec ani de zile căutând sau dezbătând lacune ale legislaţiei şi nu mai caută să amâne adoptarea măsurilor legate de transparenţă.

Ştiind că automatizarea urmează să se producă,

Mafia G ne ajută să creăm diferite scenarii pentru problema şomajului şi să ne pregătim forţa de muncă pentru cea de-a treia eră a informaticii. Cu ajutorul acestora, nu ne temem de inteligenţa artificială, ci o vedem ca pe o ocazie extraordinară pentru a spori creşterea economică şi prosperitatea fiecăruia dintre noi. Conducătorii Mafiei G reuşesc să învingă zvonurile şi să ne prezinte abordări mai eficiente pentru a ne pregăti şi educa pentru slujbele viitorului.

Strategia naţională a Statelor Unite şi formarea Coaliţiei Mafiei G inspiră şi liderii altor ţări democratice din lume să sprijine dezvoltarea globală a inteligenţei artificiale pentru binele tuturor. Universitatea Dartmouth găzduieşte ediţia inaugurală a forumului inter-guvernamental, într-o adunare care se aseamănă cu cea din vara anului 1956, unde se întâlnesc lideri din economiile cele mai avansate ale lumii: secretari, miniştri, prim-miniştri şi preşedinţi din Statele Unite, Regatul Unit, Japonia, Franţa, Canada, Germania, Italia şi alte ţări din Uniunea Europeană, precum şi cercetători din domeniul inteligenţei artificiale, sociologi, economişti, futurologi, specialişti în teoria jocului, politică şi alţii. Spre deosebire de grupul omogen de participanţi la primul atelier de la Dartmouthbărbaţi cu origini asemănătoare, de data aceasta experţii şi liderii includ un spectru larg de oameni şi perspective asupra lumii. Aflându-se pe acelaşi loc sfânt unde s-a născut inteligenţa artificială modernă, aceşti lideri acceptă să mijlocească şi să coopereze în cadrul unor iniţiative şi politici comune în domeniul inteligenţei artificiale. Inspiraţi de mitologia greacă şi de mama Pământului, ei formează GAIA: Alianţa Globală de Augmentare a Inteligenţei.

Exclusă din GAIA, China constată că influenţa sa globală este în scădere. Colaborarea internaţională nu are un impact financiar negativ asupra membrilor chinezi ai celor Nouă Titani-Baidu, Tencent şi Alibabacare continuă să ofere numeroase servicii cetăţenilor chinezi. Cu toate

acestea, multe dintre planurile pe termen lung ale Chinei printre care şi iniţiativa Noului Drum al Mătăsii sunt afectate, deoarece partenerii ies din etapa pilot şi recrutarea unor noi aliaţi se dovedeşte dificilă.

Nu sugerez că toate problemele inteligenţei artificiale dispar peste noapte. Comunitatea inteligenţei artificiale anticipează şi se aşteaptă ca inteligenţa artificială îngustă să continue să facă greşeli din cauza perspectivelor limitate asupra lumii ale membrilor iniţiali ai tribului inteligenţei artificiale. Acceptăm că prejudecăţile politice, de gen, rasă şi statut social nu vor dispărea imediat. Ţările membre GAIA semnează acorduri prin care se obligă în mod explicit să valorizeze siguranţa în defavoarea vitezei şi să aloce resurse considerabile în vederea curăţării tuturor sistemelor existente: bazele de date şi algoritmii deja implementaţi, structurile pe care acestea se sprijină, produsele pentru companii care înglobează inteligenţa artificială (precum cele folosite de bănci sau de forţele de ordine) şi produsele pentru consumatori care folosesc inteligenţa artificială pentru sarcinile zilnice (difuzoarele, ceasurile şi telefoanele noastre inteligente). GAIA pretinde – şi recompensează răspunderea în faţa publicului.

În cadrul GAIA, se ia decizia ca Dosarele de Date Personale (DDP) să fie prelucrate la fel cu registrele distribuite de blockehainuri. Registrele distribuite folosesc mii de calculatoare independente pentru a înregistra, partaja şi sincroniza tranzacţiile. Ele sunt structurate pentru a nu păstra datele centralizat, sub umbrela unei singure companii sau agenţii. Din cauză că Coaliţia Mafiei G adoptă un set de standarde şi oferă tehnologii de inteligenţă artificială unitare, DDP-urile noastre nu au nevoie de o companie coordonată central care să gestioneze tranzacţiile. Ca urmare, indivizii sunt deţinătorii propriilor DDP-uri, care sunt pe cât de publice sau de private ni le dorim şi sunt complet interoperabile-le putem conecta la unul sau toate dintre serviciile Mafiei G

simultan, dar şi în alte domenii, cum ar fi la doctor, la şcoală sau la infrastructura oraşului nostru. Membrii Mafiei G sunt custozii inteligenţei artificiale şi ai datelor noastre, însă nu şi proprietarii lor. DDP-urile sunt moştenibile: putem transmite datele noastre copiilor având capacitatea de a acorda permisiuni (de vizibilitatea completă, limitată sau zero) asupra diferitelor aspecte ale datelor proprii.

Pe măsură ce inteligenţa artificială se maturizează şi trece de la etapa aplicaţiilor înguste la maşinile gânditoare capabile de inteligenţă generală, triburile inteligenţei artificiale şi Mafia G ne-au câştigat încrederea. Nu sunt doar companii care produc aplicaţii utile – Google, Microsoft, Apple, Facebook, IBM şi Amazon sunt la fel de importante pentru America şi valorile americane ca şi baseball-ul, exprimarea liberă şi ziua de 4 iulie. Comunismul este marginalizat. Ţările care respectă dreptul cetăţenilor de a se exprima şi de a deţine proprietăţi, care sprijină libertatea religioasă, care sprijină oamenii de toate genurile, etniile, rasele şi orientările sexuale, care sunt de acord că guvernul trebuie să fie în slujba cetăţenilor, care guvernează prin intermediul unor reprezentanţi aleşi şi care pun în balanţă atât libertatea individuală şi siguranţa publică sunt aliate şi conlucrează pentru viitorul inteligenţei artificiale şi omenirii.

2029: Sugestionarea confortabilă

Graţie colaborării între membrii Mafiei G şi înfiinţării GAIA, se semnează numeroase acorduri comerciale, cetăţenii lumii au acces mai bun şi mai ieftin la produse şi servicii bazate pe IAÎ. GAIA se întruneşte în mod regulat, publicând toate dezbaterile, în timp ce grupurile sale multinaţionale de lucru ţin pasul cu progresele tehnologice.

Casele clasei de mijloc se bazează pe inteligenţa artificială pentru a-şi uşura viaţa. Aparatele, platformele şi alte servicii sunt interoperabile chiar şi între ţări, deşi în urmă cu câteva decenii licenţele şi restricţiile asupra

datelor împiedicau accesul internaţional. Maşinile de spălat şi uscătoarele folosesc mai puţină energie, sunt mai eficiente şi se conectează la sistemele oraşelor inteligente pentru a partaja date. Cu acordul nostru, permitem ca ciclurile de spălare să se efectueze atunci când solicită cel mai puţin reţelele de apă şi electricitate.

IAÎ sprijină calculele senzoriale, ceea ce înseamnă că putem strânge şi interoga lumea reală folosind date senzoriale: văzul, mirosul, auzul, gustul şi atingerea. Folosim scannere ţinute în mână, cu camere inteligente şi vizualizare digitală în bucătărie. Spectrometrul inclus într-o baghetă cu IAÎ din bucătărie citeşte lumina reflectată dintr-un avocado pentru a vă spune că nu se va coace până în weekend în timp ce uleiul de măsline pe care l-aţi cumpărat de la reduceri nu e pur ci un amestec de trei uleiuri diferite. Un alt senzor din bucătărie a detectat că puiul pus la rotisor urmează să se ardă. La etaj, un senzor tactil vă alertează că bebeluşul a scăpat (din nou) din pătuţ.

Mafia G a încheiat parteneriate cu alte companii din domeniul realităţii augmentate, îmbunătăţind semnificativ viaţa persoanelor care suferă de Alzheimer şi demenţă. Ochelarii inteligenţi recunosc automat oamenii, obiectele şi locurile, ajutându-ne rudele să-şi aducă aminte şi să aibă o viaţă mai împlinită.

Iniţial, credeam că produsele şi serviciile Mafiei G ar provoca izolare socială – că vom sta cu toţii singuri în case, interacţionând prin avataruri digitale în timp ce ne pierdem complet legătura cu lumea din afară. Ne-am înşelat complet. Platformele şi aparatura Mafiei G ne-au oferit noi moduri de a socializa pe viu. Petrecem mai mult timp în cinemauri cu realitate mixtă, care oferă un divertisment multi-senzorial. Există peste tot săli de jocuri cu realitate mixtă. E din nou ca în anii ’80, doar că s-a schimbat ceva: jocurile, experienţele şi sălile de întâlnire cu realitate mixtă sunt ieftine, şi ele pot fi folosite şi de persoanele cu tulburări de vedere şi auz. Ne ducem la

petreceri liniştite, unde purtăm căşti colorate iară fir şi ascultăm toată noaptea mixurile DJ-ului preferat. Acum cu toţii pot dansa împreună, într-o experienţă comună, chiar dacă îşi detestă reciproc gusturile muzicale. Graţie Mafiei G suntem mai conectaţi unii la ceilalţi-şi la lumea realădecât ne-am fi putut vreodată imagina.

Pentru casele mai bogate, aplicaţiile IAÎ oferă chiar şi mai multe dotări. În grădină, senzorii măsoară constant nivelul de umiditate şi compară datele cu prognozele microclimatului. Sistemele simple de irigaţii udă plantele automat, dar doar dacă este necesar. Inteligenţa artificială poate calcula nivelul optim de hidratare, ceea ce înseamnă că s-a zis cu aparatele care măsoară când trebuie udate plantele, dar şi cu begoniile moarte.

În interiorul acestor case mai bogate, sistemul Akira de la Amazon (cu o voce care nu sună nici a bărbat, nici a femeie) operează în mai multe limbi, indiferent de accent, şi poate comunica cu uşurinţă cu ochelarii inteligenţi de la Apple dar şi datele personale gestionate de Google. Maşinile de spălat şi uscătoarele sunt prevăzute cu mici drone cu braţe şi o nouă funcţionalitate denumită „modul Kondo”, în onoarea lui Marie Kondo, expertul japonez în ordonarea casei. Rufele sunt spălate şi uscate în funcţie de ciclul de cerere şi ofertă al grilei de alimentare cu curent electric, iar apoi hainele sunt preluate de o mică dronă care le împachetează şi le sortează după culori.

În Statele Unite, cumpărăturile şi livrarea lor se efectuează complet automat. Nu rămâneţi niciodată fără tampoane sau pastă de dinţi. Inteligenţa artificială stă la baza unor sistem predictive de achiziţii care îşi dau seama înaintea voastră, pe baza achiziţiilor anterioare şi a DDP-urilor voastre când trebuie să refacă stocurile. Prin intermediul Amazon puteţi cumpăra legume şi carne proaspătă din zona voastră, precum şi toate articolele de bază, cum ar fi cerealele pentru micul dejun, hârtia igienică sau chipsurile de cartofi. Serviciile de livrat truse de ingrediente la domiciliu, lansate în urmă cu un deceniu

de Blue Apron şi Hello Fresh sunt legate de PDP-ul unei case. În schimbul unei sume mai mari de bani în fiecare săptămână, cumpărăturile voastre vor include ingrediente pentru toate reţetele favorite, dar şi trei noi mese-reţete aliniate automat cu preferinţele, antipatiile, alergiile şi nevoile nutritive ale fiecărui membru al familiei.

Veţi continua desigur să cumpăraţi şi în lumea reală, dar veţi putea alege să vă lăsaţi portofelul acasă. Tehnologia care stă la baza sistemelor comerciale Amazon Go a devenit coloana vertebrală a magazinelor cu servire rapidă în care o bună parte a inventarului este deja la vedere sau poate fi uşor înlocuit. Camerele inteligente supraveghează constant cumpărătorii, recunoscând amprentele faciale şi conţinutul coşurilor de cumpărături. Putem cheltui până la 100 de dolari fără să fim nevoiţi să interacţionăm cu un angajat. În magazinele cu suprafeţe mai mari (de ex., magazine universale, de mobilă sau de bricolaj) sau cele care vând marfa mai scumpă (de ex., bijuterii, genţi şi electronice) avem opţiunea să plătim cu faţa.

Unii copii se joacă cu animale vii, în timp ce familiile cu un stil de viaţă mai agitat optează pentru animale robotizate. Mici câini şi pisici – nişte receptacule drăguţe ale inteligenţei artificialefolosesc calculele senzoriale şi învăţarea profundă în timp ce sunt dresaţi să respecte regulile casei. Dispunând de camere avansate în fosele oculare, de blană cu senzori tactili şi de capacitatea de a recunoaşte modificările subtile ale vocilor noastre, animalele robotizate sunt mult mai empatice decât cele vii, chiar dacă sunt mai puţin calde şi pufoase.

Oricine, indiferent de nivelul de venit, se bucură de faptul că i se sugestionează un stil de viaţă mai sănătos. Mafia G ne aduce aminte în timpul zilei să facem alegeri mai sănătoase. În timp ce vă îndreptaţi spre muncă şi aşteptaţi liftul, ceasul vibrează un pic pentru a vă face să vă uitaţi în jos: vă arată o hartă simplă a clădirii de birouri, iar o săgeată vă îndreaptă către scări. Cu siguranţă că

puteţi dezactiva această opţiune, însă majoritatea oamenilor preferă să o lase pornită. Şi antrenamentele personale sunt optimizate. Folosind dosarul vostru de date personale, istoricul medical şi datele senzoriale strânse din multe alte sursecăştile iară fir prin care ascultaţi muzică, ţesătura inteligentă din care este croit sutienul sportivechipamentul de fitness vă ghidează prin exerciţii personalizate. După ce aţi terminat antrenamentul, senzorii vă ajută să vă răcoriţi, monitorizându-vă metabolismul şi inima. Datorită Mafiei G, comunităţile noastre sunt mai sănătoase, iar vieţile noastre sunt mai lungi.

Prin faptul că Mafia G a agreat un singur standard pentru DDP-uri, interfeţele utilizatorilor, formatul de arhivare a datelor, protocoalele şi structurile folosite în sistemul medical sunt standardizate. Ca urmare, sistemul de sănătate este mult mai eficient. Politicienii s-au certat preţ de decenii referitor la asistenţa sanitară în America, iar insistenţa Mafiei G pe algoritmi şi date standardizate în sănătate s-a dovedit a fi cel mai bun leac.

Indiferent ce doctor vede un pacient, sau la ce spital este internat, datele acestuia pot fi accesate de orice persoană responsabilă cu îngrijirea lui. Ele sunt disponibile şi pentru orice persoană căreia pacientul i-a acordat permisiunea. Datele din majoritatea testelor de laborator, analizelor şi scanărilor sunt analizate de inteligenţa artificială şi nu de oameni, conducând la o precizie mai mare şi la procesarea mai rapidă a rezultatelor. Sistemul IBM poate detecta anomalii celulare pentru a depista semnele timpurii ale cancerului – precum şi care sunt celulele afectate din corp. Sistemul Google ajută doctorii să prevadă rezultatul probabil al diferitelor medicamente şi tratamente, precum şi să estimeze când va muri un pacient, ajutând personalul medical să ia decizii mai bune în privinţa tratamentului pentru fiecare pacient. Aflat în spital, un pacient primeşte medicamentele acasă ca urmare a sincronizării între farmacia Amazon şi DDP.

Chiar şi dacă în istoricul medical al pacientului sunt pagini întregi de reţete scrise de mână de către doctorişi chiar dacă aceste observaţii clinice nu sunt foarte amănunţitesistemul de vizualizare digitală şi analiză a tiparelor al Mafiei G completează aceste detalii, transformând aceste informaţii în date structurate, care pot fi folosite şi extrase doar pentru pacient, sau anonimizate şi combinate cu alte date despre pacienţi pentru a sprijini comunitatea medicală (formată atât din oameni cât şi din inteligenţe artificiale) în a-şi spori cunoaşterea şi experienţa.

Diagnosticul, tratamentul şi îngrijirea nu sunt oferite doar în spitale fizice, ceea ce înseamnă că mult mai mulţi oameni din Statele Unite au acum un acces mai bun la sănătate. Deşi relativ recente, unii furnizori oferă acum servicii de medicină la distanţă. Toaletele TOTO, dotate cu spectrofotometru şi eprubete pot folosi detectarea tiparelor pentru a diagnostica un nivel crescut sau scăzut de glucoză sau proteine, precum şi bacteriile sau hematiile. După câteva secunde, DDP reflectă o posibilă infecţie urinară sau semnele incipiente ale pietrelor la rinichi. Tratamentele simplecum ar fi antibioticele pentru infecţii sunt comparate cu DDP-ul vostru, trimise medicului de familie, iar dacă acesta aprobă, sunt livrate automat la voi acasă, la muncă sau în oraş la cină.

Periuţele de dinţi, dotate cu mici senzori de salivă analizează lichidul pentru a vedea care este starea generală de sănătate. Cu fiecare periaj, inteligenţa artificială monitorizează nivelul hormonilor, electroliţilor şi anticorpilor, verificând dacă acesta este constant de-a lungul timpului. Mafia G a schimbat standardele de îngrijire: testele de diagnosticare elementară nu sunt doar pentru pacienţii bolnavi, ci fac parte din menţinerea unui stil de viaţă sănătos. Această evoluţie a schimbat la rândul ei caracterul medicinei, ea trecând de la o abordare reactivă la una predictivă, profilactică.

Şi alte aspecte ale vieţii cotidiene printre care şi

viaţa amoroasă sau sexulsunt mai bune graţie inteligenţei artificiale. Algoritmii evolutivi s-au dovedit o soluţie mai inteligentă pentru persoanele care-şi caută dragostea pe Internet decât aplicaţiile şi paginile de Internet existente până atunci. Cercetătorii au stabilit că oamenii sunt pur şi simplu prea complecşi pentru a putea fi reduşi la un set de date pe care un singur algoritm trebuie să le compatibilizeze. Pe lângă acest aspect, avem tendinţa să completăm profilurile online oferind informaţii care descriu mai degrabă cum ne-am dori să fim – şi nu cum suntem de fapt. Algoritmii evolutivi preiau datele din DDP-uri şi ne verifică toate profilurile din bazele de date. Ne alegem un obiectiv – de la „vreau doar să mă distrez” la „sunt gata să mă căsătoresc” apoi adăugăm parametrii de limitare (trebuie să fie de religie mozaică, şi să trăiască la maximum 70 de km de Cleveland), iar algoritmul evolutiv generează o listă de oameni cu care avem cele mai mari şanse să ne atingem acel obiectiv. Dacă ne dorim, sistemul va consulta agenda şi preferinţele de activităţi şi va programa automat data şi locul întâlnirii. După câteva întâlniri (sau dacă prima întâlnire nu a mers prea bine), ne-ar putea interesa să folosim un algoritm generativ care să creeze conţinut pornografic personalizat. În funcţie de preferinţele noastre, inteligenţa artificială creează scene care să ne excite, să ne inspire sau să ne înveţe, cu personaje ale căror voci, corpuri şi stiluri se mulează pe dorinţele noastre.

Graţie Mafiei G, inteligenţa artificială nu ne lasă impresia că ar înlocui creativitatea omenească, ci că o completează e un instrument prin care să ne mărim şi sporim inteligenţa. În cadrul firmelor de arhitectură, inteligenţa artificială generează mii de posibile clădiri pe baza modelelor preferate de client şi a constrângerilor precum şi alege şi ierarhizează planurile câştigătoare pe baza predicţiilor privind fezabilitatea luând în calcul termenul de execuţie, materialele disponibile şi bugetul; dificultatea obţinerii autorizaţiilor necesare şi dacă

clădirea afectează negativ traficul pietonal. Investitorii imobiliari folosesc inteligenţa artificială pentru a simula durabilitatea pe termen lung, dat fiind climatul unei zone dar şi alţi factori de mediu. Meseriaşii calificaţi – tâmplari, electricieni şi instalatori-folosesc ochelari cu realitate mixtă oferiţi de Google, Microsoft şi o companie pe nume Magic Leap pentru a vedea prin pereţi, a verifica dacă proiectul se potriveşte cu execuţia şi a detecta de la bun început eventualele probleme.

Folosirea creativă a inteligenţei artificiale se manifestă şi în arte, inclusiv în cinematografie. Se sărbătoresc 20 de ani de la premiera filmului *Avatar*, pelicula care în 2009 părea dintr-o altă lume graţie efectelor speciale hiper-realiste, generate pe calculator. Pentru a sărbători această aniversare, Cameron dezvăluie un proiect secret bazat pe inteligenţă artificială: un al şaselea film *Avatar* bazat pe o tehnologie de detectare a mişcării subacvatice pe care o dezvoltase anterior, precum şi un nou mediu virtual şi un sistem de proiecţie pe retină. Experienţa este construită folosind algoritmi de generare meniţi să creeze lumi complet noi pe care să le exploreze avatarurile umane, algoritmi evolutivi pentru randări şi învăţare profundă pentru efectuarea tuturor calculelor necesare. Rezultă un film unic prezentat într-un context special care (împreună cu sistemul de proiectare pe retină) oferă o experienţă narativă complet originală-şi care introduce spectatorul complet în subiect.

Inteligenţa artificială ajută organizaţiile de orice dimensiune să aibă abordări de management mai creative. Mafia G oferă modele predictive pentru analiza de afaceri, oferind modalităţi de eficientizare, reducerea costurilor şi domenii care ar putea fi îmbunătăţite. Departamentele de resurse umane folosesc recunoaşterea tiparelor pentru a evalua productivitatea şi moralul angajaţilor şi pentru a elimina prejudecăţile la angajare sau promovare. Nu mai există CV-uri; DDP-urile noastre ne arată punctele forte şi slăbiciunile, iar programele de inteligenţă artificială ne

analizează datele înainte de a ne recomanda unor posibili angajatori.

În cadrul multor companii mari, muncitorii umani nu mai prestează activităţi cu nivel cognitiv scăzut, în timp ce inteligenţa artificială sprijină angajaţii în anumite domenii legate de cunoaştere. Activitatea recepţionerilor, a asistenţilor de clienţi, a agenţilor de turism este acum pe deplin automatizată. În cadrul întâlnirilor, difuzoarele inteligente ascultă, folosind algoritmii de detectare a vocii pentru a transcrie conversaţiile noastre. O secretară virtuală ia notiţe automat, subliniind numele vorbitorilor, punctele importante discutate, zonele de convergenţă şi divergenţă, informaţii contextuale din alte întâlniri şi alte date relevante ale companiei. Sistemul stabileşte temele care trebuie monitorizate ulterior şi creează liste de activităţi pentru participanţi.

Pentru că am recunoscut încă de la bun început că automatizarea va bulversa anumite părţi ale forţei de muncă, şomajul nu este răspândit iar economia noastră are fundaţii puternice. În Statele Unite, guvernul federal pune la dispoziţie programe de reconversie profesională pentru a se asigura că suntem bine pregătiţi. Folosind instrumentele Mafiei G, companiile şi persoanele s-au pregătit din timp pentru noile tipuri de slujbe.

Mafia G a pus la dispoziţia şcolilor publice şi private inteligenţa artificială pentru a spori procesul de învăţare. Sistemele de învăţare capabile de adaptare, supravegheate de profesori, îi fac pe elevi să înveţe în propriul ritm, în special în ce priveşte cititul, logica, matematica şi limbile străine. În sălile de clasă şi în case, IBM l-a reînviat pe filosoful Socrate sub forma unui agent de inteligenţă artificială cu care dialogăm şi care ne pune întrebări şi ne răspunde pentru a ne stimula gândirea critică. Sistemul de inteligenţă artificială Socrate, evoluat din Watson, pune întrebări studenţilor legate de materie, dezbătând şi discutând conceptele. (Inteligenţa artificială Socrate este folosită şi în afara şcolii, fiind un membru apreciat al

oricărei echipe medicale, juridice, de ordine publică, strategie şi politică. Ea este folosită şi pentru a pregăti candidaţii pentru dezbaterile publice.)

Inteligenţa artificială Socrate a IBM e un aliat util în redacţiile de ştiri, ajutând jurnaliştii să analizeze ştirile şi posibile perspective alternative. Ea este folosită şi pentru verificarea datelor şi asigurarea calităţii editoriale: articolele sunt verificate pentru a nu conţine prejudecăţi inconştiente şi a se asigura că se foloseşte un amestec variat de surse şi perspective. (Au apus de mult timp vremurile în care ziarele şi revistele publicau liste doar cu bărbaţi gânditori eminenţi, oameni de afaceri de succes şi aşa mai departe.) Algoritmii de generare sunt folosiţi pentru a crea videoclipuri din cadre fotografice, a modela tridimensional peisaje şi clădiri pe baza câtorva fotografii şi a detecta voci distincte ascunse în mulţimi. Prin urmare apar mult mai multe reportaje video care sunt mai ieftin de produs.

Inteligenţa artificială este folosită pentru a detecta tipare şi anomalii ale datelor, făcând jurnaliştii să scoată la iveală noi relatări în interesul public. În loc să fie un complice al roboţilor de dezinformare, inteligenţa artificială poate detecta propaganda, afirmaţiile false şi campaniile de dezinformare. Ca urmare, democraţiile sunt mai puternice.

Mafia G a studiat oraşele din China în care se desfăşoară proiecte pilot ale oraşelor inteligente – cum ar fi Rongeheng, Beijing, Shenzhen, Shanghaişi a identificat cele mai bune practici care ar putea fi preluate şi în Statele Unite. Avem acum câteva oraşe inteligente în America-Baltimore, Detroit, Boulder şi Indianapoliscare testează o gamă largă de servicii şi sisteme de inteligenţă artificială. Reţelele de sateliţi cubici-mici sateliţi de mărimea unui cub Rubiktransferă date în timp real sistemelor de inteligenţă artificială care pot recunoaşte obiecte, tipare unice de lumină şi amprente termice. Aceasta le permite edililor să prevadă eventuale pene de

curent, să monitorizeze şi redirecţioneze traficul, să gestioneze rezervele de apă şi să deszăpezească drumurile. Inteligenţa artificială îi ajută şi să gestioneze bugetele şi personalul pe parcursul anului, descoperind modalităţi complet noi de a economisi la scară largă. Cheltuielile neprevăzute nu dispar, dar impactul lor se diminuează-iar cetăţenii acestor oraşe sunt animaţi de o senzaţie de speranţă pe care nu au mai trăit-o de mulţi ani.

Aceste sisteme furnizează date şi autorităţilor de ordine publică, precum poliţia sau pompierii, care folosesc inteligenţa artificială pentru a analiza cantităţi uriaşe de date, inclusiv fluxuri video: dacă nu există sunet, algoritmii de recunoaştere a tiparelor pot citi de pe buze şi pot transcrie conversaţiile. Algoritmii de generare completează şi ei lipsurile datelor audio iar dacă o înregistrare este neclară, un algoritm o poate face mai clară. Inteligenţa artificială scanează milioane de imagini, căutând tipare pe care un ochi omenesc le-ar putea rata. Instalarea acestor sisteme nu s-a făcut fără controverse. Cu toate acestea, angajamentele Mafiei G în domeniul protecţiei intimităţii fac ca DDP-urile noastre să nu poată fi accesate fără mandat. Avem siguranţa că Mafia G ne apără intimitatea.

Pe măsură ce evoluează, inteligenţa artificială ne ajută să devenim nişte oameni mai buni. Prin rolul activ jucat de Mafia G, guvernul federal şi GAIA, ne simţim confortabil cu noţiunea de sugestionare.

2049: Rolling Stones au murit (dar fac muzică nouă)

Până în anii ’2030, cercetătorii care lucrează în cadrul Mafiei G au publicat un articol stimulant, atât prin prisma a ceea ne dezvăluie despre inteligenţa artificială, dar şi a felului în care a fost scris. Pornind de la acelaşi set de standarde şi având sprijinul unei finanţări (şi răbdări) substanţiale a guvernului federal, cercetătorii au colaborat pentru evoluarea inteligenţei artificiale. Ca urmare, a fost dezvoltat primul sistem capabil de inteligenţă artificială generală.

Sistemul a trecut testul Colegului. A durat mult timp ca în comunitate să se accepte faptul că testul Turing, şi alte teste similare, nu reprezintă barometrul corect pentru măsura inteligenţa maşinilor. Testele care se bazează fie pe amăgire (poate un computer să păcălească un om că discută că un om?) sau pe imitare (poate un computer să se comporte la fel cum am face-o noi?) nu recunosc particularitatea inteligenţei artificiale: o inteligenţă obţinută şi exprimată într-o manieră care nu se aseamănă cu experienţa umană. În loc să judecăm o IAG în funcţie de capacitatea acesteia de a „gândi” la fel cum o facem noi, comunitatea inteligenţei artificiale a adoptat în sfârşit un nou test pentru a măsura *contribuţiile semnificative* ale unei LAG, care să judece valoarea sarcinilor cognitive şi comportamentalediferite, dar puternicepe care nu le-am putea efectua pe cont propriu. LAG se consideră atinsă atunci când sistemul poate contribui în mod general de o manieră similară sau superioară unui om.

Mafia G a cercetat şi dezvoltat timp de ani de zile o LAG care să poată asista la o şedinţă de lucru şi să contribuie cu o sugestie valoroasă-fară ca să i se ceară acest lucruânainte de finalul şedinţei. Această IAG a primit numele de cod „Proiectul Hermione”, după personajul din *Harry Potter* care ştia întotdeauna şi în orice situaţie ce să spună sau să facă. Majoritatea oamenilor de pe glob au adus contribuţii valoroase în cadrul unui grup, la un moment dat: la muncă, într-un context religios, la barul din cartier cu prietenii sau într-un curs de istorie din liceu. Simpla întrerupere doar de dragul de a prezenta o informaţie sau de a răspunde la o întrebare nu sporeşte valoarea unei conversaţii. O contribuţie însemnată presupune o combinaţie a mai multor aptitudini:

Capacitatea de a face presupuneri informate: Ea se mai numeşte şi „raţionament pe furiş” şi reprezintă felul prin care reuşim să ne descurcăm în viaţă. Folosim cele mai bune informaţii pe care le avem, enunţăm şi verificăm ipoteze şi găsim răspunsuri chiar şi în absenţa unor

explicaţii clare.

Extragerea corectă a semnificaţiei cuvintelor, pauzelor şi zgomotului de fond: Doar pentru că cineva spune că e „încântat” să se ocupe de un nou proiect nu înseamnă că e *încântat* la propriu. Alte indicii, precum limbajul lor corporal, ne-ar putea transmite că sunt destul de *neîncântaţi* să accepte cererea, însă nu sunt în stare să spună nu.

Folosirea experienţei, cunoaşterii şi a contextului istoric pentru a înţelege: Atunci când oamenii interacţionează, ei aduc cu sine o perspectivă nuanţată asupra lumii, un set unic de experienţe personale şi, de regulă, şi propriile aşteptări. Uneori, logica şi datele nu sunt suficiente pentru a câştiga o dezbatere. Alteori, ele sunt singurele aspecte care contează. Interpretarea încăperii: Există o interacţiune explicită-şi una tacită, care se petrece în profunzime. Indiciile subtile ne pot ajuta să ne dăm seama atunci când este o problemă considerabilă care trebuie să ne capteze atenţia.

Proiectul Hermione a participat la o sesiune a unui grup de lucru al GAIA. Optsprezece membri ai grupului au discutat şi dezbătut standardele existente pentru inteligenţa artificială, dezvoltate fie de către oamenii aflaţi în încăpere, fie de predecesorii acestora. Deoarece grupul era divers şi alcătuit din lideri din diferite ţări şi culturi, există multe informaţii subliminale: anumite dinamici de putere, conflicte de personalitate şi senzaţii de inferioritate sau superioritate. Grupul a discutat cu IAG ca un membru egal, fără beneficii suplimentare sau excepţii speciale. Pe la jumătatea sesiunii, IAG a respins un consens uşor majoritar, dar cu potenţial de creştere, care se degajase în jurul temei reglementărilor. IAG a argumentat cu tact împotriva ideii şi a obţinut sprijinul unui alt membru al grupului pentru a susţine o alternativă. Proiectul Hermione a făcut o contribuţie importantă. (De nepreţuit, ar putea susţine alţii ulterior.)

Proiectul Hermione a fost un succes nu doar pentru

că a reuşit să treacă testul Colegului cu atâta uşurinţă-ci şi din cauză că GAIA şi Mafia G au văzut în acest moment atât un semnal de alarmă cât şi o oportunitate. Aceste organizaţii au continuat să-şi recalibreze strategiile şi standardele pentru a putea ţine pasul cu evoluţia tehnologică a inteligenţei artificiale. Organizaţiile au decis să limiteze ritmul auto-îmbunătăţirii, adăugând constrângeri tuturor sistemelor bazate pe inteligenţă artificială pentru ca oamenii să fie tot timpul informaţi. Cercetătorii GAIA urmează acum noi protocoale: ei rulează simulări pentru a putea înţelege impactul unor LAG mai puternice înainte de a le permite să fie folosite în scopuri generale, comerciale sau militare.

Membrii Mafiei G sunt companii bogate, puternice şi cu influenţă-iar succesul lor sporeşte. Aceste companii construiesc aplicaţii practice interesante ale IAG pentru a ne spori productivitatea şi creativitatea, contribuind şi la oferirea de soluţii pentru problema cea mai presantă a omenirii: schimbările climatice. Pe măsură ce curentul Atlantic s-a deplasat înspre nord, şi grânarul Americii a urmat acelaşi drum, înspre Canada, decimând fermele şi agricultura Statelor Unite. Cafeaua şi ciocolata nu mai pot fi crescute cu uşurinţă în aer liber. Cetăţenii din Bangladesh, Filipine, Thailanda şi Indonezia au devenit refugiaţi climatici în propriile ţări. Amazon, în parteneriat cu Microsoft, grupul francez Danone şi DowDuPont din Statele Unite folosesc IAG şi experimentele genetice pentru a creşte legume în sere.

Google şi Facebook folosesc LAG pentru a muta întregi populaţii în siguranţă, creând noi comunităţi omeneşti. IAG îi ajută să găsească locurile în care pot trăi cel mai uşor o viaţă confortabilă şi care să păstreze cultura popoarelor afectate. Regiunile altă dată pustii ale planetei sunt modelate sau transformate folosind materiale de construcţii adaptabile. Zgârie-pământurilecomplexe mari, întinse, de doar câteva etaje înălţimemodifică complet amprenta urbană. În interior, lifturi fără cabluri ne

transportă în orice direcţie. E o tendinţă nouă în arhitectură care ajută cele mai importante centre economice ale lumii să prospere – inclusiv oraşele americane Denver, Minneapolis şi Nashville.

O perioadă părea că China se va retrage şi se va regrupa în jurul puţinilor aliaţi rămaşi-Coreea de Nord, Rusia, Mongolia, Myanmar, Cambodgia, Kazahstan, Pakistan, Kârgâzistan, Ţadjikistan şi Uzbekistan. Universităţile din ţările GAIA nu mai acceptă studenţi din China. Din frica de supraveghere şi pierdere a DDP-urilor, industria turismului dispare din China. Ţările GAIA se bazează pe sisteme automate care să producă materiile prime pentru fabricaţie, reluând fabricarea produselor în ţară. În cele din urmă, guvernul chinez a stabilit că excluderea din GAIA era nocivă pentru economie, provocând tulburări politice şi economice considerabile. China a acceptat cu o oarecare reticenţă să adopte normele şi standardele GAIA, precum şi măsurile de transparenţă cerute de statele membre GAIA. Comunismul nu a muritexistă încă destule puncte politice fierbinţi, dar şi tensiunile obişnuite legate de diferitele stiluri de guvernare şi conducere.

LAG nu a apărut fără multe probleme noi, unele putând fi anticipate. Ca şi alte tehnologii care au transformat cu timpul societatea omenească, LAG a eliminat slujbe, a condus la apariţia unor noi tipuri de activităţi ilicite şi a scos la iveală partea urâtă a noastră. În anii ’2040, LAG nu este însă o ameninţare existenţială.

Folosim LAG atât la muncă cât şi acasă pentru a accesa informaţii. LAG este un agent de control care ia multiple forme, în funcţie de situaţie: vorbim cu ea, interacţionăm cu ea pe un ecran şi îi trimitem date provenite din corpurile noastre. Toate familiile au un valet, pentru că în fiecare casă există o LAG antrenată şi calibrată în funcţie de împrejurările casei.

Una dintre cele mai mari şi mai notabile schimbări aduse de LAG este o sporire drastică a sofisticării în

majoritatea aspectelor existenţei umane. Putem să fim recunoscători Mafiei G pentru creşterea considerabilă a calităţii vieţii. Activităţi dificile, cronofage – cum ar fi găsirea unui timp convenabil pentru toată lumea, stabilirea unui program de activităţi după şcoală, gestionarea finanţelor personale sunt acum complet automatizate şi supravegheate de către LAG. Nu ne mai pierdem ore întregi încercând să facem curat în căsuţa poştală-LAG ne înlesneşte sarcinile cognitive facile. Avem în sfârşit aparate electrocasnice care ne curăţă podelele şi covoarele, ne pun rufele la uscat şi ne şterg praful. (Ne aducem aminte de 2019 ca de o epocă mult mai simplă, plină de sarcini manuale monotone, plictisitoare.)

Răceala nu mai există, şi nici gripa. De fapt, ne mirăm acum cât de naivi erau doctorii din trecut. Şi asta pentru că IAG ale IBM şi Google ne-au ajutat să vedem şi înţelegem milioane de viruşi diferiţi. Acum, când nu ne simţim bine, o IAG efectuează un test de diagnostic pentru a stabili cauza astfel încât să poată fi prescris un tratament în funcţie de DDP. Au dispărut şi medicamentele fără reţetă, însă preparatele magistrale cunosc o revenire. Aceasta se datorează faptului că LAG a contribuit la accelerarea unor evoluţii critice în domeniul modificărilor genetice şi al medicinei de precizie. Puteţi apela acum la serviciile unui farmacist informatic: farmaciştii special pregătiţi cu pregătire în bioinformatică, medicină şi farmacologie.

Farmacia informatizată este o specializare medicală separată, care colaborează îndeaproape cu un nou tip de medici de familie cu formare în inteligenţă artificială: medici generalişti cu studii de medicină şi tehnologie. Deşi LAG a eliminat anumiţi specialişti: radiologii, imunologii, alergologii, cardiologii, dermatologii, endocrinologii, anesteziştii, neurologii şi alţii doctorii care activau în aceste domenii au avut suficient timp la dispoziţie pentru a dobândi noi competenţe în domenii adiacente. Viaţa pacientului este mai bună. Nu mai petreceţi ore în şir de la

un doctor la altul, primind mesaje contradictorii şi nu vi se mai prescriu medicamente în exces. Dacă trăiţi într-o zonă mai izolată, apariţia IAG v-a îmbunătăţit semnificativ accesul la îngrijirea medicală.

De la naştere, fiecărui copil i se scanează genomul acest procedeu este rapid şi la îndemâna oricui. Oamenii preferă să aleagă această variantă pentru că secvenţa genetică reprezintă o componentă esenţială a DDP-ului. Pe lângă faptul că vă oferă o fereastră prin care puteţi privi alcătuirea voastră genetică unică, IAG analizează datele pentru a detecta anomalii genetice şi a afla mai multe despre cum funcţionează corpul vostru. Desigur, atât în Statele Unite cât şi în alteţări există în continuare minorităţi care se opun acestei practiciaşa cum existau cândva oameni care se opuneau vaccinării. Deşi părinţii pot alege să se opună din motive religioase sau ideologice, puţini aleg această cale.

Graţie IAG suntem mai sănătoşi-şi avem la dispoziţie noi opţiuni când vine vorba de căutarea unui partener şi de căsătorie. Forme avansate de intimitate diferenţiată permit unui terţ să acceseze datele (DDP-ul, genomul şi istoricul medical) fără ca să ştie cui aparţin aceste date. Furnizorii de servicii matrimoniale care au implementat tehnologia IAG devin astfel deosebit de utili, pentru că acum putem alege să optimizăm în funcţie de viitoarea familie (copii din combinaţii genetice dezirabile), avere (estimarea potenţialului de câştig de-a lungul vieţii) sau distracţie (dacă vor râde sau nu la glumele voastre).

LAG vă ajută şi în alte demersuri creative, dincolo de căutarea iubirii. Membrii Rolling Stones au murit cu mulţi ani în urmă, dar graţiei algoritmilor mimetici, ei continuă să creeze muzică. Aceeaşi senzaţie pe care aţi avut-o când aţi auzit prima dată primele 30 de secunde ale melodiei „Paint It Black” – melodia melancolică de chitară, urmată de 8 bătăi puternice de tobe şi un refren repetitiv în care Mick Jagger cântă „I see a red door and I want it painted black”. – un moment unic de stimulare şi satisfacţie. Nu

părea posibil că veţi putea simţi acelaşi lucru din nou faţă de un cântec al Rolling Stones, şi cu toate acestea cea mai nouă melodie a lor este la fel de zgomotoasă, de dură şi de stimulantă.

Deşi ziarele pe hârtie au dispărut, presa a adoptat LAG ca metodă de distribuţie. Odată ce LAG au putut trece testul Colegului, redacţiile au reacţionat rapid pentru a construi un nou model de distribuţie a ştirilor, care să poată fin inc continuare profitabil dar săţmă mai mult cont de preferinţele viitorului. În zilele noastre, majoritatea oamenilor nu „primesc” ştirile sau nu „dau drumul” la ştiriei au o discuţie cu un agent de ştiri inteligent. *New York Times* şi *Wall Street Journal* angajează sute de jurnalişti informatici-oameni cu pregătiri hibride în inteligenţa artificială şi jurnalism. împreună, aceste echipe relatează ştiri şi aleg datele relevante care trebuie incluse în motoarele de conversaţie. Jurnalismul bazat pe IAG ne informează, şi putem să-l configurăm pentru a-i da o culoare politică, a primi mai multe date secundare sau o variantă extinsă cu personaje secundare şi fapte diverse. Suntem consultaţi în ce priveşte analiza ştirilor şi a fluxurilor editoriale, dezbătând în mod constructiv cu agentul de ştiri folosind vocea sau interacţionând cu ecranele (ochelari inteligenţi şi tablete retractabile). Există în continuare articole de amploare, atât sub formă text cât şi video.

Hackerii IAG în mare parte *alte IAG* sunt o problemă constantă pentru că aceste IAG comit ilegalităţi, dezvăluind oamenii care le-au creat codul sursă. Agenţiile guvernamentale angajează ofiţeri cu experienţă în analiza datelor. Cu ajutorul grupului BAT din China, cei Nouă Titani colaborează la dezvoltarea unei aparaturi, a unor structuri, reţele şi algoritmi capabile să reziste atacurile. Parteneriatul dintre GAIA şi INTERPOL a reuşit în mare măsură să limiteze criminalitatea organizată.

Proiectele pilot de oraşe inteligente lansate în urmă cu două decenii în Baltimore, Indianapolis, Detroit şi

Boulder au fost de succes şi au ajutat şi alte comunităţi să-şi dezvolte bunele practici, conducând la înfiinţarea Administraţiei Federale a Infrastructurii Inteligente. Ca şi Administraţia Federală a Autostrăzilor, AFII operează sub egida Departamentului Transporturilor şi supraveghează toate sistemele conectate care ne alimentează oraşele: staţii de transformare fără cabluri, generatoare descentralizate de energie (cinetice, solare şi eoliene), reţele de conectare a vehiculelor la infrastructură, fibra optică prin care se aduce lumina soarelui în fermele subterane. Datele senzoriale sunt adunate şi folosite pentru a modela sănătatea generală a comunităţilor noastre: accesul la un aer curat, curăţenia cartierelor, folosirea parcurilor şi a zonelor exterioare de agrement. IAG prevede şi elimină penele de curent şi penuriile de apă înainte de a se produce.

Pe măsură ce ne apropiem de tranziţia de la IAG la SIA, se întrevede o oportunitate extraordinară: interfeţele între creier şi maşină. Ne aflăm pe pragul de a deţine nanotehnologia moleculară, sperând ca în câteva decenii să putem înregistra simultan datele generate de miliardele de neuroni din creierul nostru. Computere microscopice, de dimensiunea unui fir de nisip, ar urma să se stea în vârful creierului şi să capteze semnalele electrice. Sisteme IAG speciale, capabile să citească şi să interpreteze aceste semnale, ar putea de asemenea să transmită date între oameni. O interfaţă creier-maşină ar putea cândva permite unei persoane sănătoase să redea funcţionalitatea unei persoane care a suferit un accident vascular cerebral, o paralizie sau care nu mai poate vorbi. Interfeţele creier-maşină, pe care le-am putea folosi teoretic pentru a transfera amintiri între oameni, ne-ar putea ajuta să trăim empatia de o manieră mai profundă şi mai plină de semnificaţie.

Acest potenţial ne face să ne gândim la noi utilizări ale LAG. Vrem să aflăm răspunsul la întrebări filosofice spinoase: *Trăim într-un univers real*? *E posibil ca „nimicul”*

să existe? *Care e natura timpului*? LAG nu ne poate da răspunsurile pe care ni le dorim, însă Mafia G ne poate spori înţelegerea condiţiei umane.

2069: Gardienii Galaxiei cu puterea inteligenţei artificiale

Explozia inteligenţei, anticipată în urmă cu 100 de ani de matematicianul şi pionierul inteligenţei artificiale I.J. Good, începe la finalul anilor ’2060. Devine clar că LAG atinge niveluri profunde de inteligenţă, viteză şi putere, şi că super-inteligenţa artificială poate fi atinsă în viitorul apropiat. În ultimul deceniu, cei Nouă Titani şi GAIA s-au pregătit pentru acest evenimentealculând că odată ce nivelul inteligenţei omeneşti va fi depăşit de maşini, SLA ar putea să apară după doar câţiva ani.

După multe dezbateri, membrii GAIA iau decizia dificilă de a împiedica apariţia unei SLA. Unii dintre cei care au participat s-au emoţionat, argumentând că nu e just ca „minţile strălucite” ale inteligenţei artificiale să fie mutilate taman când se apropiau de potenţial. Dezbatem dacă negăm sau nu omenirii posibilitatea de a avea câştiguri şi oportunităţi chiar mai mari.

În cele din urmă, cu binecuvântarea şi încurajarea celor Nouă Titani, GAIA stabileşte că toate LAG trebuie să aibă restricţii care să le limiteze rata auto-îmbunătăţirii şi care să împiedice apariţia unor mutaţii nedorite, pentru a apăra siguranţa şi securitatea oamenilor. În curând, GAIA va lansa o serie de paznici bazaţi pe inteligenţă artificială, care să acţioneze ca un sistem de avertizare timpurie pentru orice LAG care capătă prea multă putere cognitivă. Chiar şi gardienii nu vor împiedica o entitate ilicită să încerce să creeze propriile SLA, însă GALA scrie scenarii pentru a putea contracara o astfel de situaţie. GALA şi cei Nouă Titani beneficiază de afecţiunea şi încrederea noastră nelimitate.

CAPITOLUL ŞASE

CUM SĂ ÎNVĂŢĂM SĂ TRĂIM CU MILIOANE DE TĂIETURI CU HÂRTIA: SCENARIUL PRAGMATIC

Până în 2023, problemele inteligenţei artificiale sunt cunoscute dar decidem pe parcurs să modificăm doar în mică măsură parcursul de dezvoltare al inteligenţei artificiale, un sistem despre care ştim cu toţii că prezintă defecte evidente. Urmărim doar mici retuşuri pentru că cei care controlează inteligenţa artificială nu sunt dispuşi să-şi asume măsuri incomode: să renunţe la potenţialele beneficii financiare, să ia decizii politice nepopulare, şi să ne tempereze aşteptările exagerate pe termen scurt, chiar dacă asta ar însemna îmbunătăţirea şanselor de a trăi pe termen lung alături de inteligenţa artificială. Chiar mai grav este faptul că ignorăm China şi planurile sale de viitor.

Liderii din Congres, din diversele agenţii federale şi de la Casa Albă continuă să nu acorde o prioritate prea mare inteligenţei artificiale şi cercetării ştiinţifice avansate în general, preferând să investească în industrii cu mai mare priză din punct de vedere politic dar care sunt aproape de finalul vieţii. Planul strategic privind viitorul inteligenţei artificiale publicat de administraţia Obama în 2016 – un document care a influenţat puternic planul strategic chinez pentru anul 2025 – este abandonat, împreună cu programul federal de finanţare a cercetării şi dezvoltării în domeniul inteligenţei recomandat de acest plan. America nu are o viziune sau o strategie pe termen lung în domeniul inteligenţei artificiale, şi nu crede că aceasta va avea vreun impact economic, educaţional sau asupra securităţii naţionale. Politicienii americani, din ambele partide, se concentrează asupra înăbuşirii Chinei, în loc să dezvolte strategii pentru a coaliza forţele Mafiei G şi ale guvernului. Absenţa unei coaliţii şi a unei strategii naţionale coerente în privinţa inteligenţei artificiale generează tăieturi cu hârtiamilioane şi milioanecare încep să sângereze pe parcurs. Ele nu se observă de la început. Cultura pop, articolele jurnaliştilor din domeniul tehnologiei şi postările de pe reţelele de socializare ale persoanelor influente ne-au învăţat să căutăm semne mari,

evidentecum ar fi roboţi ucigaşiprin urmare, vom rata adevăratele semnale de alarmă, mai mici şi mai puţin evidente, pe măsură ce inteligenţa artificială evoluează. Cei Nouă Titani sunt forţaţi să acorde prioritate vitezei în detrimentul siguranţei, astfel încât parcursul de dezvoltare al inteligenţei artificiale – de la L-AI către IAG şi mai departecontinuă fără să rezolve mai întâi vulnerabilităţile tehnice grave. Iată câteva dintre tăieturile cu hârtia mai puţin evidentemulte fiind auto-indusepe care nu le tratăm ca nişte răni serioase, aşa cum vor ajunge să devină.

În calitate de consumatori de tehnologie, avem aşteptarea ca triburile inteligenţei artificiale să-şi fi imaginat şi să fi rezolvat toate problemele înainte ca aplicaţiile, produsele sau serviciile să părăsească laboratoarele de cercetare şi dezvoltare. Ne-am obişnuit să adoptăm o tehnologie care funcţionează imediat ce o scoatem din cutie. Când cumpărăm noi telefoane inteligente şi televizoare, le băgăm în priză, iar ele funcţionează aşa cum ni se promite. Când descărcăm programe noi, fie că sunt procesoare de text sau programe de analiză a datelor, ele se comportă conform aşteptărilor. Uităm că inteligenţa artificială nu este o tehnologie care să funcţioneze imediat ce o scoatem din cutie, pentru că pentru a putea funcţiona aşa cum ne dorim, un sistem de inteligenţă artificială are nevoie de cantităţi uriaşe de date şi de ocazia de învăţa în timp real.

Niciunul dintre noi-consumatori, jurnalişti sau analiştinu permite celor Nouă Titani să facă greşeli. Pretindem noi produse, servicii, invenţii sau descoperiri, şi ne lamentăm în public când nu le primim. Nu ne interesează că solicitările noastre distrag atenţia triburilor inteligenţei artificiale de la perfecţionarea produselor.

Modelele şi structurile inteligenţei artificiale, indiferent cât de mari sau de mici, au nevoie de multe date pentru a putea învăţa, a se îmbunătăţi şi a putea fi implementate. Datele sunt ca oceanele lumii. Ne înconjoară, constituie o resursă nelimitată, şi complet

inutilă dacă nu o desalinizăm, tratăm şi procesăm pentru a o putea consuma. În momentul de faţă există doar câteva companii care pot desaliniza eficient datele la o scară care să conteze. Acesta este motivul pentru care partea cea mai dificilă a construirii unui nou sistem de inteligenţă artificială nu se rezumă la algoritmi sau la modele, ci la colectarea datelor potrivite şi etichetarea lor corectă pentru ca maşina să se poată antrena şi învăţa de pe urma lor. Comparativ cu diversitatea produselor şi serviciilor pe care cei Nouă Titani se străduie în grabă să le construiască, există foarte puţine seturi de date care sunt gata de a fi folosite. Câteva dintre acestea sunt Image Net (un set uriaş de imagini folosit la scară largă), Wiki Text (un set de date pentru modelarea limbajului bazat pe articole din Wikipedia), 2000 HUB5 (un set de date disponibil doar în engleză folosit pentru vorbire) şi Libri-Speech (aproximativ 500 de ore de mostre din audiobook-uri. Dacă aţi dori să construiţi o inteligenţă artificială pentru sănătate care să detecteze anomalii în analizele de sânge sau scanările oncologice, problema nu e inteligenţa artificială, ci lipsa dateloroamenii au organisme complicate, cu multe variaţii, şi nu există un set suficient de mare de date gata de a fi folosite.

În urmă cu un deceniu, în anii ’2010, echipa IBM Watson Health a încheiat un parteneriat cu mai multe spitale pentru a vedea dacă inteligenţa artificială ar putea sprijini activitatea doctorilor. Watson Health avea la activ nişte realizări remarcabile, inclusiv un caz în care era implicat un băieţel bolnav în vârstă de 9 ani. După ce specialiştii nu au putut să-l diagnosticheze şi să-l trateze, Watson a atribuit o probabilitate mai multor probleme de sănătatepe listă fiind atât boli mai frecvente cât şi mai exotice, inclusiv un sindrom rar întâlnit în copilări denumit boala Kawasaki. Odată ce s-a auzit că Watson pune diagnostice miraculoase şi salvează vieţi omeneşti, echipa Watson a fost supusă unor presiuni de a comercializa platforma, fiind stabilite ţinte nerealiste. IBM proiectase că

Watson Health avea să crească de la venituri de 244 de milioane de dolari în 2015, la 5 miliarde de dolari până în 2020.175 Aceasta ar fi reprezentat o creştere de 1949% în mai puţin de cinci ani.

Înainte ca Watson Health să poată să stârnească din nou aceeaşi senzaţie ca în trecut-respectând în acelaşi timp şi un program extrem de ambiţios de dezvoltaresistemul ar fi avut nevoie mult mai multe date şi de timp pentru a învăţa. Nu existau însă suficiente date de sănătate din lumea reală, iar cele existente nu erau suficient de amănunţite. Motivul era că aceste date erau stocate în arhive gestionate de o altă companie, care considera că IBM îi face concurenţă.

Ca urmare, echipa IBM a folosit o variantă ocolitoare frecventă întâlnită în rândul triburilor inteligenţei artificiale. Sistemul Watson Health a fost alimentat cu aşa-zise „date sintetice”, respectiv date care reprezintă informaţii ipotetice. Deoarece cercetătorii nu dispun de date brute pe care să le introducă într-un sistem de învăţare pentru a-l pregăti, ei cumpără un set de date sintetice de la o terţă firmă sau construiesc chiar ei unul. Această abordare este adesea problematică pentru că alcătuirea acelui set de datece date trebuie incluse, cum trebuie etichetatedepinde de deciziile unui număr redus de oameni, care adesea nu sunt conştienţi de propriile prejudecăţi politice, profesionale, de gen, şi altele.

Aşteptările exagerate vizavi de profitabilitatea imediată a Watson Health, combinate cu o dependenţă de seturile de date sintetice au condus la o problemă gravă. IBM încheiase un parteneriat cu Memorial Sloan Kettering Cancer Center pentru a folosi Watson Health în domeniul oncologic. La scurt timp după aceea, câţiva experţi care colaborau la proiect au relatat exemple de recomandări incorecte sau hazardate de tratament. Într-un astfel de caz, Watson Health a recomandat un protocol bizar de tratament unui pacient diagnosticat cu cancer pulmonar care manifesta şi semne de hemoragie: chimioterapie şi un

medicament denumit bevacizumab, care ar fi fost contraindicat pentru că poate provoca hemoragii grave sau chiar fatale. 176 Vestea că Watson e inept s-a răspândit în publicaţiile de specialitate din domeniul sănătăţii şi pe blogurile informaticienilor, fiind adesea acompaniată de titluri incendiare. Cu toate acestea, problema nu consta în faptul că Watson Health voia să le facă oamenilor de hac, ci că forţele de piaţă au exercitat presiuni asupra IBM pentru a grăbi ritmul cercetării în inteligenţa artificială pentru a îndeplini planurile financiare.

Iată o altă tăietură cu hârtia: unele inteligenţe artificiale şi-au dat seama cum să trişeze şi să-şi modifice propriile sisteme. Dacă o inteligenţă artificială este programată cu scopul de a învăţa un joc, a-l juca şi a face orice e nevoie pentru a câştiga, cercetătorii au descoperit cazuri în care recompensele au fost manipulate, prin faptul că un sistem se va folosi de algoritmi evolutivi şi de învăţare pentru a câştiga folosind înşelătoria şi disimularea. Spre exemplu, o inteligenţă artificială care joacă Tetris şi-a dat seama că ar putea pur şi simplu să pună jocul pe pauză pe termen nelimitat pentru a nu pierde niciodată. Încă de când aţi citit prima dată despre manipularea recompenselor – ştirea a devenit celebră recent când două sisteme de inteligenţă artificială din domeniul financiar au prezis o scădere dramatică a bursei şi au încercat să închidă pe cont propriu bursele pe termen nelimitatv-aţi întrebat ce s-ar putea întâmpla dacă datele voastre ar fi exploatate într-un sistem de manipulare a recompenselor. Concediul acele din vacanţa de iarnă care urmează cum ar fi dacă serviciile de control al traficului aerian nu ar mai funcţiona?

O altă tăietură cu hârtia: agenţii răuvoitori pot introduce date otrăvite în programele de pregătire a inteligenţei artificiale. Reţelele neurale sunt vulnerabile la „exemplele proaste” care sunt informaţii fictive sau eronate menite să inducă intenţionat un sistem de inteligenţă artificială să facă o greşeală. Un sistem de

inteligenţă artificială ar putea eticheta o imagine ca fiind un panda, cu un grad de siguranţă de 60%, dar dacă adăugăm un pic de distorsiuni imaginii, câţiva pixeli pe care un om nu i-ar detecta, iar sistemul ar putea re-eticheta imaginea ca un gibon, cu o siguranţă de 99%. E posibil să păcălim sistemul de vizualizare digitală al unei maşini pentru a crede că un semafor roşu transmite de fapt „limită de viteză 100”, iar pasagerii să treacă la viteză maximă printr-o intersecţie. Datele eronate ar putea reprograma un sistem militar de inteligenţă artificială să interpreteze semnele unui spital obişnuit – ambulanţe, cuvântul „urgenţă” sau „spital” ca aparţinând unor organizaţii teroriste. Problema e că cei Nouă Titani nu şi-au dat seama cum ar putea să-şi protejeze sistemele de exemplele proaste, nici în lumea digitală, nici în cea fizică.

O tăietură mai profundă: cei Nouă Titani ştiu că informaţiile false pot fi folosite pentru a reprograma sistemele automate de învăţare şi reţelele neurale. O echipă din cadrul diviziei Brain a Google a publicat în 2018 un articol prin care explica cum un agent malefic ar putea introduce informaţii false într-o bază de date dedicată vizualizării digitale şi să reprogrameze practic tot ce învaţă sistemele din acea bază de date.177 Hackerii ar putea într-o bună zi să introducă date otrăvitoare în căştile voastre inteligente şi să le reprogrameze cu identitatea altei persoane doar prin redarea unor sunete nocive în preajma voastră în metrou.

Situaţia se complică prin faptul că informaţiile contrare pot avea şi o utilitate. O altă echipă a Google Brain a descoperit că informaţiile contrare pot fi folosite într-o „reţea antagonică generativă” sau RAG, pe scurt. E ca un test Turing, dar fără niciun om. Două inteligenţe artificiale sunt programate folosind aceleaşi datenişte imagini cu oameni, de pildă. Prima generează fotografii ale dictatorului nord coreean Kim Jong-un care să pară realiste, în timp ce a doua compară fotografiile generate cu cele reale. Pe baza analizei celei de-a doua inteligenţe

artificiale, prima revine şi rafinează propriul proces. Secvenţa se repetă până când prima inteligenţa artificială generează automat tot soiul de imagini cu Kim Jong-un care par complet realiste, dar care nu au existat niciodată în lumea reală. Imagini care-l arată pe Kim Jong-un luând cina cu Vladimir Putin, jucând golf cu Bernie Sanders sau sorbind cocktailuri cu Kendrick Lamar. Google Brain nu urmăreşte subterfugiile, ci să rezolve problema datelor sintetice. RAG ar permite sistemelor de inteligenţă artificială să lucreze cu date brute, din lumea reală, care nu au fost filtrate şi fără supravegherea directă a unui programator uman. Şi deşi e o abordare deosebit de creativă pentru rezolvarea unei probleme, ea ar putea într-o zi să constituie o ameninţare gravă la adresa siguranţei noastre.

Încă o tăietură cu hârtia: când algoritmii complecşi colaborează, ei concurează uneori unul împotriva celuilalt pentru a îndeplini un obiectiv, iar acest fapt poate otrăvi un sistem întreg. Am fost martorii unor probleme sistemice atunci când preţul unui manual de biologie a început să crească galopant. Cartea nu se mai tipărea, dar pe pagina Amazon se afişa că există 15 copii folosite de la alţi comercianţi, începând de la un preţ de 35, 54 dolarişi două copii noi nouţe la un preţ de 1, 7 milioane de dolari. În spatele site-ului, algoritmii Amazon au demarat un război al preţurilor, alegând să ridice preţul din ce în ce mai mult până când a ajuns la 23.698.655 dolari (plus 3 dolari şi 99 de cenţi pentru expediere). Sistemul algoritmilor de învăţare ajustase în timp real ca reacţie la fiecare licitaţie, iar aceste operaţii făceau parte din bagajul cu care fusese programat. Ca să reformulăm: e posibil să fi învăţat fără să vrem inteligenţa artificială că bulele de pe pieţe sunt un lucru bun. Nu e greu să ne imaginăm algoritmi rivali care umflă preţul caselor, acţiunilor sau chiar al publicităţii digitale până la cote absurde.

Acestea sunt doar o mică parte dintre tăieturile cu hârtia pe care triburile inteligenţei artificiale au hotărât că

le putem suporta de dragul atingerii obiectivelor stabilite de forţele de pe piaţă în Statele Unite, respectiv de Partidul Comunist Chinez la Beijing. În loc să domolească pretenţiile legate de viteză şi profitabilitate, triburile inteligenţei artificiale sunt presate în continuu să scoată produse pe piaţă. Siguranţa e doar un aspect secundar. Angajaţii şi conducerea

Mafiei G sunt îngrijoraţi, însă nu le permitem suficient timp pentru a face schimbările necesare. Şi încă nu am vorbit deloc de China.

Între 2019 şi 2023 am ignorat efectiv dezideratele lui Xi Jinping pentru viitor: strategia naţională amănunţită a Chinei în domeniul inteligenţei artificiale, planurile lui de a domina economia globală, şi obiectivul Chinei de a deveni o forţă unică de decizie în domeniul geopolitic. Am eşuat în a lega viitorul inteligenţei artificiale de infrastructura de supraveghere şi sistemul de credit sociale, precum şi activitatea diplomatică chineză în diferite ţări africane, asiatice şi europene. Aşa că, atunci când Xi vorbeşte public şi frecvent despre nevoia unei reforme a sistemului politic mondial şi continuă înfiinţând entităţi internaţionale precum Banca Asiatică de Investiţii în Infrastructură, nu îi acordăm toată atenţia, ci doar o privire cu coada ochiului. Aceasta este o greşeală pe care nu o recunoaştem imediat.

În China, parcursul către o poziţie dominantă în inteligenţa artificială nu a fost lipsit de obstacole. China are propriile tăieturi cu hârtia cu care se confruntă, deoarece grupul BAT se luptă să poată inova în acelaşi ritm cu Silicon Valley sub mâna forte a Beijingului. Grupul BAT eludează în mod repetat reglementările birocratice. Toate aceste scandarluricum ar fi cel prin care Administraţia de Stat a Rezervelor Valutare din China a amendat Alipay cu suma de 600.000 de yuani (aproximativ 88.000 de dolari) pentru înregistrarea contabilă eronată a unor plăţi internaţionale între 2014 şi 2016, în timp ce Tenpay a fost pedepsită pentru că nu a depus documentele corecte de înregistrare în vederea efectuării de plăţi

internaţionale între 2015 şi 2017-s-au dovedit a nu reprezenta doar excepţii.178 Devine evident că acestea nu reprezintă incidente izolate iar oficialii chinezi trebuie să se confrunte cu tensiunile dintre abordările socialiste şi realităţile capitalismului.

Vedem deja consecinţele tuturor acestor vulnerabilităţi de ordin politic, strategic şi tehnic. Pentru a face pe plac finanţatorilor, Mafia G preferă să încheie contracte profitabile cu guvernul în locul unor parteneriate strategice. Astfel, se stimulează un climat concurenţial în locul colaborării. Consecinţa este o interoperabilitate redusă a structurilor, serviciilor şi aparatelor care folosesc inteligenţa artificială. La începutul anilor ’2020, piaţa a sugestionat Mafia G să-şi împartă anumite funcţionalităţi şi activităţi: Amazon domină în casele noastre şi în comerţul electronic, în timp ce Google domină în domeniul căutărilor pe internet, al localizării, comunicării personale şi la locul de muncă. Microsoft domină în domeniul cloud, în timp ce IBM şi-a arogat aplicaţiile inteligenţei artificiale la nivel de companii şi sisteme aplicate de sănătate. Reţele de socializare sunt dominate de Facebook, iar Apple fabrică aparate (telefoane, computere, accesorii).

Niciun membru al Mafiei G nu aderă la un singur set de valori de bază care să acorde prioritate transparenţei, incluziunii şi siguranţei. Deşi conducătorii companiilor din Mafia G sunt de acord că ar trebui probabil ca în domeniul inteligenţei artificiale să existe standarde adoptate la scară largă şi ca ele să fie implementate, nu există pur şi simplu o modalitate prin care să se aloce timp sau resurse pentru ele.

Dosarul de Date Personale este construit, întreţinut şi deţinut de patru companii ale Mafiei G: Google, Amazon, Apple şi Facebook. Iată însă şi problema: nici măcar nu ştiţi că aceste DDP-uri există sau că Mafia G şi triburile inteligenţei artificiale se folosesc de ele. Nu este o omisiune intenţionată, ea datorându-se pur şi simplu vitezei. Existenţa lor se regăseşte în termenii şi condiţiile

pe care le acceptăm întotdeauna, fără să le citim niciodată.

Formatul folosit de fiecare furnizor de DDP nu este complementar aşa că există multe date duplicat care există în mai multe locuri, dar şi, în mod paradoxal, multe spaţii goale unde lipsesc date. E ca şi cum patru fotografi diferiţi v-ar face o poză: una cu reflectoare şi umbrele de reflecţie, una cu un obiectiv ochi-depeşte, una cu o cameră de unică folosinţă şi una cu un aparat RMN. Din punct de vedere tehnic, sunt patru imagini ale capului vostru, dar datele pe care le conţin cele patru sunt radical diferite.

Străduindu-se să obţină o imagine mai completă, triburile inteligenţei artificiale trimit „ambasadori digitali” – mici programe care acţionează ca intermediari şi negociază în numele Mafiei G. Ambasadorii digitali de la Google şi Amazon funcţionează o perioadă, dar nu reprezintă soluţii realiste pe termen lung. Aceştia sunt prea dificil de actualizat, mai ales că atât de multe produse şi servicii oferite de terţi sunt legate de ei. În loc să lanseze variante noi zilnic, Google alege să facă o mare schimbare.

La începutul anilor ’2020, Google lansează penultimul sistem de operare, un mega sistem de operare care poate rula pe telefoane inteligente, difuzoare inteligente, laptopuri, tablete şi electrocasnice cu funcţionalităţi legate de internet. Acesta este doar începutul. În final, Google urmăreşte să crească şi să îmbunătăţească acest sistem de operare astfel încât să devină infrastructura invizibilă care ne ajută în viaţa de zi cu zi, operând interfeţele, căştile şi ochelarii inteligenţi, maşinile şi chiar anumite componente ale oraşelor. Acest sistem este complet integrat cu DDP-urile şi aduce îmbunătăţiri semnificative tuturor utilizatorilor săi.

Mega sistemul de operare al Google este lansat într-o perioadă nefastă pentru Apple, care a devenit prima companie americană care a atins o capitalizare de piaţă de o mie de miliarde de dolari, dar care înregistrează un declin constant al vânzărilor de telefoane iPhone pe fondul

unui val de noi aparate conectate, precum căştile sau ceasurile inteligente. În ciuda multelor sale reuşite, Amazon (a doua companie americană care a atins capitalizarea de o mie de miliarde de dolari) nu a mai avut o reuşită în domeniul electronicelor de la difuzorul inteligent Echo încoace. Într-o mutare surpriză, Apple şi Amazon încheie în 2025 un parteneriat pentru a dezvolta un sistem de operare care să ruleze pe aparatele produse de ambele companii. Sistemul de operare rezultat-Applezoneconstituie o ameninţare formidabilă la adresa Google. Pentru consumatori, această evoluţie instaurează un model bazat pe două sisteme de operare şi dă startul unei consolidări masive şi rapide a ecosistemului inteligenţei artificiale.

Facebook decide că trebuie să încheie un parteneriat similar; numărul utilizatorilor activi este în scădere, iar aceştia nu mai consideră reţeaua de socializare ca fiind indispensabilă. Facebook încearcă să se apropie de Applezon, care nu se arată interesaţi. Microsoft şi IBM continuă să se axeze pe serviciile oferite companiilor.

China şi noii săi parteneri diplomatici folosesc împreună tehnologiile BAT, în timp ce restul lumii foloseşte fie mega sistemul de operare al Google, fie Applezon, ambele sisteme folosind şi dând valoare DDP-urilor noastre. Piaţa ne oferă alegeri limitate. Există doar câteva modele de telefoane (iar ochelarii şi brăţările inteligente vor înlocui în curând telefoanele mobile) dar şi pentru toate aparatele din casele noastre: difuzoare, computere, televizoare, electrocasnice sau imprimante. E mai uşor să alegem o singură marcă, astfel există case Google sau case Applezon. Din punct de vedere tehnic, ne-am putea transfera DDP-urile la alţi furnizori, însă nu deţinem personal datele din DDP-uri şi nici DDP-urile însele. Nu beneficiem de o transparenţă totalănu ştim în bună măsură ce fac Google şi Applezon cu DDP-urile noastre pentru că prin felul în care şi-au creat sistemele au urmărit protejarea proprietăţii intelectuale.

Pentru a evita procesele anti-trust, ni se spune că putem porta oricând DDP-urile între sistemele de operare. În practică e însă aproape imposibil să faceţi tranziţia. Vă aduceţi aminte cât de greu era acum mulţi ani când încercaţi să treceţi de la iOS la Android şi descopereaţi că multe date şi setări importante se pierdeau pentru totdeauna, că nu se păstra activitatea din cadrul aplicaţiilor, multe dintre ele nici măcar nu funcţionau (şi nu puteaţi obţine rambursarea banilor plătiţi pentru ele) şi că toate locurile în care vă depozitaţi înainte pozele şi videourile nu mai puteau fi accesate cu uşurinţă. Acum că DDP-ul vostru este folosit şi de terţi-şcoli, spitale sau linii aeriene, trecerea de la Google la Applezon sau viceversa este mult mai dificilă.

Există mulţi consultanţi în informatică specializaţi în portarea DDP-urilor de la un furnizor la celălalt, însă procesul durează câteva zile, este costisitor şi imperfect. Majoritatea oamenilor aleg cu reticenţă să rămână unde sunt, chiar dacă rezultatele nu sunt optime.

Google şi parteneriatul Amazon-Apple se confruntă cu procese anti-trust atât în Statele Unite cât şi în Europa. În timp ce procesele se târăsc în instanţă, toate datele sunt atât de amestecate încât separarea sau deschiderea sistemelor DDP şi ale inteligenţei artificiale ar genera riscuri mai mari decât recompensele potenţiale. Ca urmare, se ia decizia ca cele două companii să fii amendate cu sume considerabilefondurile urmează să fie folosite pentru dezvoltarea unor noi companii. Cu toţii sunt însă de acord că sistemul cu doi furnizori de sisteme de operare trebuie să continue.

În timp ce inteligenţa artificială se maturizează şi trece de la aplicaţii limitate la maşini capabile de inteligenţă generală, nu avem de ales şi trebuie să trăim cu tăieturile cu hârtia provocate de inteligenţa artificială. Varianta modernă a comunismului chinezsocialism amestecat cu nuanţe de capitalismise extinde, permiţându-i lui Xi Jinping să îndeplinească promisiunile legate de

noua ordine mondială. Ţările care se opun stilului autocratic chinezesc, suprimării presei libere şi libertăţilor religioase şi perspectivelor negative asupra anumitor etnii, orientări sexuale sau de gen nu deţin suficientă putere. Ele nu au de ales şi trebuie să colaboreze cu China, în termenii impuşi de aceasta.

Ni s-a promis că inteligenţa artificială ne va aduce libertatea, eliminând sarcinile plictisitoare şi munca repetitivă. În schimb, libertatea noastră de a alege este restricţionată de o manieră pe care nimeni nu şi-a imaginat-o.

2029: Neajutorarea învăţată

Regimul binar de sisteme de operare a generat o concurenţă acerbă între membrii triburilor inteligenţei artificiale care nu au anticipat problemele mari de interoperabilitate. Se pare că pe lângă aparate, nici *oamenii* nu sunt interoperabili în acest regim. Lejeritatea odinioară emblematică a Silicon Valleyingineri, directori de operaţiuni sau designeri ai experienţei utilizatorilor obişnuiau să treacă de la o companie la alta fără niciun simţ adevărat de apartenenţă a dispărut de mult timp. În loc să ne aducă împreună, inteligenţa artificială a reuşit efectiv şi eficient să ne izoleze. Şi pentru guvernul Statelor Unite reprezintă o problemă, fiind forţa să aleagă o structură. (Ca majoritatea celorlalte guverne şi cel american a ales Applezon în defavoarea Google, pentru că Applezon oferea preţuri mai mici şi rechizite la preţ redus).

În toată lumea se vorbeşte despre „neajutorarea învăţată” în epoca inteligenţei artificiale. Părem că nu putem funcţiona fără diversele sisteme automate, care ne sugestionează în mod constant cu feedback negativ sau pozitiv. Încercăm să dăm vina pe cei Nouă Titani, dar de fapt doar noi suntem de vină.

Pentru membrii generaţiei Y, lovitura a fost deosebit de grea, deoarece ei erau însetaţi să primească feedback şi laude când erau copii şi ei au iubit la început diversele sisteme ale inteligenţei artificiale, însă între timp au

căpătat un tic psihologic care e greu de eliminat. Când se termină bateria unei periuţe de dinţi, un membru al acestei generaţii (care are acum peste 40 de ani) trebuie să-şi perieze dinţii cu o periuţă obişnuită, care nu oferă niciun feedback Neprimind niciun feedback de confirmare, persoana nu primeşte dopamina la care se aşteaptă, simţindu-se anxioasă şi tristă. Fenomenul nu se întâlneşte doar la membrii acestei generaţii. O senzaţie de nelinişte ne bântuie pe toţi. Investim în lucruri redundante, ne cumpărăm obiecte analogice de rezervă (cum ar fi periuţele de dinţi din plastic, căştile vechi sau ochelarii Warby Parker) pentru a le completa pe cele care conţin inteligenţă artificială. Ne-am pierdut încrederea în fostul bun simţ şi în aptitudinile de bază necesare vieţii.

Standardele rivale ale Google şi Applezon ne aduc aminte de ce înseamnă să călătoreşti în străinătate şi toate necazurile cu prize de format şi tensiune diferite. Cei care călătoresc în mod regulat acordă prioritate alegerilor în funcţie de sistemul de operare şi nu de programele de fidelitate, alegând un hotel Applezon sau o linie aeriană Google. Companiile constată că e mai uşor să aleagă doar unul dintre cele două sisteme. Încet dar sigur, suntem sugestionaţi să alegem una dintre tabere. Cei care preferă Applezon au dificultăţi în a trăi cu cei care preferă Google, pentru că DDP-urile şi aparatele lor nu sunt compatibilechiar dacă personalităţile lor sunt.

Anul 2019 a marcat începutul sfârşitului telefoanelor inteligente, motiv pentru care acum purtăm accesorii conectate şi nu le mai ţinem în buzunare sau în genţi. După o perioadă de progrese rapide, noile telefoane care rulau iOS sau Android ofereau doar îmbunătăţire minore ale sistemelor, în vreme ce pentru telefoane nu existau îmbunătăţiri majore în afara fotografiei. Entuziasmul din jurul flecării lansări a unor noi modele iPhone s-a pierdut. Nici măcar lansarea unui telefon Samsung cu ecran flexibil nu a fost suficientă pentru a convinge consumatorii să-şi abandoneze modelele mai vechi. În loc să stea la coadă

după un an sau doi pentru a achiziţiona cel mai nou telefon, consumatorii au ales să cheltuie acei bani pe noi aparate: căşti Bluetooth fără fir cu senzori biometrici, brăţări care vă permit să înregistraţi clipuri video şi să efectuaţi apeluri video, şi ochelari inteligenţi care ne oferă un flux aparent infinit de informaţii. Applezon a reuşit să lanseze prima pe piaţă ochelarii Applezon Visionşi nu a fost nicio surpriză. Apple şi Amazon aveau un istoric lung şi plin de reuşite în promovarea noilor tehnologii şi modelarea preferinţele consumatorilor. (Eşecul comercial al Google Glass continuă să fie dureros pentru unii angajaţi ai companiei, chiar dacă tehnologia era revoluţionară.) în zilele noastre majoritatea oamenilor poarte ochelari şi căşti inteligente toată ziua, cu o brăţară sau un inel cu care pot efectua înregistrări video.

Se pare că ochelarii erau de neevitat. După două decenii în care ne-am holbat în ecrane, ochii noştri nu mai pot face ajustările necesare şi majoritatea dintre noi avem o vedere înceţoşată la distanţă şi avem nevoie de ochelari de vedere de la vârste mai fragede. Ca majoritatea oamenilor şi voi aveţi nevoie de ochelari, creând o piaţă pentru ochelarii inteligenţi pe care unii analişti o credeau imposibilă. Ochelarii, împreună cu căştile fără fir, o brăţară inteligentă şi o tabletă uşoară, reprezintă principalul mijloc de comunicare. Vă uitaţi la videouri prin ei, iar pentru a efectua un apel video folosiţi camera din brăţara inteligentă. În general, vorbiţi mai mult decât tastaţi. Algoritmi speciali pentru procesare în spaţiu, vizualizare digitală şi recunoaştere vocală stau la baza majorităţii datelor pe care le vedeţi şi le înregistraţi prin intermediul acestor accesorii inteligente.

Applezon şi Google v-au sugestionat să închiriaţişi nu să deţineţiaceste echipamente, iar accesul la DDP face parte din abonament. Nu e nimic necurat în acest model al abonamentului, decizia fiind luată din considerente de ordin practic legate de ciclul de viaţă al produselor. Rata evoluţiei în domeniul inteligenţei artificiale accelerează în

fiecare an, iar de vreme ce valoarea datelor noastre este semnificativ mai mare decât marja de profit a ochelarilor, brăţărilor şi căştilor inteligente, obiectivul este de a ne menţine pe toţi conectaţi la sistemul. Tehnologia se vinde în pierdere, fiind compensată de un abonament lunar accesibil. Pe baza acestui abonament obţineţi şi accesul la DDP, care este evaluat în funcţie de permisiunile acordate companiilor. Abonamentele cele mai ieftine oferă şi anonimitatea cea mai scăzută, prin urmare clienţii care aleg această variantă acordă libertatea Google şi Applezon să folosească datele lor personale după bunul plac, fie că e vorba de publicitate sau de simularea unor experimente medicale. Persoanele mai întărite pot adăuga „prime de permisiune” pachetelor DDP, dar acestea se obţin cu mare dificultate şi costă foarte mult. În 2029, există comunităţi ale elitelor, cu pază sporită şi departe de ochii lumii- acestea sunt însă comunităţi digitale, păzite de algoritmi şi care protejează datele persoanelor bogate de ochii curioşi ai oamenilor şi companiilor obişnuite.

Ca mulţi alţii, aţi fost victime ale unor aşa-zise „atacuri ale papagalilor”, o nouă variantă de înşelătorii digitale, iar guvernele din lumea întreagă au fost luate complet prin surprindere. Se pare că programele ilegale vă pot infecta şi DDP-ul, imitând ca un papagal vocea voastră către toţi cunoscuţii. Unele inteligenţe artificiale de tip papagal sunt atât de bine înrădăcinate în dosarele voastre şi în viaţa voastră digitală încât pot imita convingător nu doar vocea, ritmul, tonul şi vocabularul vostru, dar o pot face cu date reale din viaţa voastră. Aceste programe sunt folosite pentru a trimite mesaje vocale false pe care mulţi părinţi sau soţi le cred adevărate. Din nefericire, aceste programe de inteligenţă artificială papagal provoacă probleme companiilor care oferă servicii matrimoniale pe internet. Escrocii fură identitatea altora şi o folosesc pentru a atrage oamenii folosind interacţiuni hiper-realiste.

Neajutorarea învăţată, noile bariere economice şi

senzaţia că sinele nostru real nu poate concura cu varianta îmbunătăţită prin inteligenţa artificială ne provoacă multe amărăciuni. Căutăm consolarea prin interfeţele creier-maşină, legături cu capacitate mare care transferă date între creier şi un calculator. Deşi Facebook şi Elon Musk au anunţat acum un deceniu că lucrează la un aparat special care ne-ar da super-puteri telepatice, Baidu a fost prima companie care a lansat „bentiţa neurostimulatoare”. Ascunsă într-o şapcă, ea poate citi şi monitoriza undele cerebrale şi transmite un feedback pentru a creşte concentrarea, a crea o senzaţie de fericire şi mulţumire, sau a vă face să credeţi că aveţi multă energie. Nu a fost de mirare că o companie a grupului BAT a reuşit să dezvolte prima o astfel de interfaţă. Companiile farmaceutice au făcut lobby pe lângă autorităţi, în speranţa că acestea vor bloca autorizarea bentiţelor neurostimulatoare şi a unor interfeţe viitoare între creier şi maşină. Considerând că Baidu e o ameninţare.

Google şi Applezon au intervenit, lansând propriile produse, ceea ce a făcut ca DDP-urile noastre să conţină şi mai multe date.

Sâcâiala e noua sugestionare, iar Google şi Applezon vă hărţuiesc neintenţionat pentru a avea o sănătate mai bună. Brăţara, căştile şi ochelarii inteligenţi vă comunică în mod constant mesaje de reamintire. Nu apucaţi să mâncaţi o linguriţă de tort, pentru că din clipa în care vă uitaţi către desert, inteligenţa artificială recunoaşte ce urmează să mâncaţi, compară mâncarea cu ritmul metabolic şi starea de sănătate actuală şi vă trimite o notificare de avertizare către brăţară sau ochelari. La restaurant, vi se sugestionează să alegeţi mâncărurile care se potrivesc cu nevoile biologice actuale: alimente cu conţinut ridicat de potasiu sau omega-3, sau alimente cu conţinut scăzut de carbohidraţi sau sare. Dacă alegeţi inteligent, sunteţi recompensaţi şi primiţi mesaje de încurajare.

Nu aveţi nicio variantă reală de a vă deconecta de la

inteligenţa artificială sâcâitoare, deoarece DDP-ul este legat de prima de asigurare de viaţă, calculată pe baza angajamentului de a duce o viaţă sănătoasă. Dacă săriţi peste o sesiune recomandată de antrenament, vă puteţi aştepta să fiţi sâcâit toată ziua. Dacă mâncaţi o prăjitură în plus, fapta e consemnată la dosar. Sistemul nu a fost creat să se comporte în acest fel, însă algoritmilor li s-a oferit un scop şi au fost învăţaţi să optimizeze fără întrerupere diferitele aspecte ale vieţii cotidiene. Nu au fost programaţi cu un punct de final sau o dată de încetare.

Când s-a instaurat regimul cu două sisteme de operare pentru DDP-urile noastre, mulţi dintre furnizorii de arhive medicale au trebuit să aleagă un partener. Unii membri ai Mafiei G au obţinut astfel datele de care au avut nevoie în urmă cu mai mulţi ani, creândântrucâtva întâmplător-noul sistem sanitar american. IBM Watson Health deţinea tehnologia complexă (unii ar putea susţine chiar superioară) dar avea şi un istoric de disfuncţionalităţi organizaţionale de două decenii. La 15 ani după ce Google a lansat Calico, propriul proiect de sănătate, acesta nu a reuşit să scoată pe piaţă produse viabile din punct de vedere comercial, astfel încât un parteneriat strategic părea soluţia: Watson Calico. Din partea Google a fost o mişcare inspirată deoarece atât Amazon cât şi Apple plănuiau propriile iniţiative în domeniul asigurărilor şi farmaceuticelor. Amazon experimentase cu noi modele de asigurare şi livrare a medicamentelor printr-un parteneriat cu Berkshire Hathaway şi J.P. Morgan, în timp ce Apple şi-a folosit lanţul de magazine şi modelul Genius Bar pentru a lansa o nouă marcă de clinici la minut pe toată coasta de vest a Statelor Unite. Parteneriatul dintre Google şi IBM a forţat o a doua colaborare de tip Applezon, între platforma electronică a farmaciilor Amazon şi clinicile la minut ale Apple. Ca urmare a acestor consolidări de pe piaţă, spitalele din America fac astăzi parte fie din sistemul Watson-Calico, fie din Applezon Health Marile companii de asigurări-Kaiser Permanente, Life Point Health, Trinity

Health, New York-Presbyterian Healthcare Sistem sunt înscrise fie în reţeaua Watson-Calico, fie în Applezon Health.

Aceste colaborări s-au dovedit o soluţie extraordinară la problema datelor. Google, IBM şi Applezon au acum acces nelimitat chiar şi la datele voastre biologice-iar voi primiţi la schimb acces la diagnostice gratuite sau cu cost scăzut. Analizele nu se efectuează doar când sunteţi bolnavi. Sunteţi analizaţi pentru orice, lucru care are un efect benefic asupra nivelului de sănătate şi fericire. Dacă întrebaţi orice american ce temperatură are corpul lui, veţi primi un răspuns personalizat, şi nu cele 37, 5 grade Celsius standard.

Deşi avem în sfârşit acces la asistenţă sanitară accesibilă, americanii trebuie acum să suporte nişte defecte bizare care s-au dovedit a fi trăsături ale sistemului şi nu erori. Ambulanţele mai vechi sau asistentele medicale din şcoli sau din taberele de vară nu pot accesa întotdeauna DDP-ul pacientului dacă nu au cele mai recente actualizări de sistem. DDP-urile sistemelor rivale de sănătate pot fi citite din punct de vedere tehnic atât de Applezon Health cât şi de Watson-Calico, însă adesea lipsesc multe date contextuale utile, în special în comunităţile mai mici sau cele rurale, doctorii constată că trebuie să-şi aducă aminte ce au învăţat în facultate dacă o persoană care locuieşte într-o casă Applezon se prezintă la o clinică Watson-Calico, sau invers. Pe măsură ce doctorii instruiţi în metodele mai vechi, tradiţionale se pensionează, există din ce în ce mai puţini doctori tineri care au cunoştinţele şi experienţa necesară pentru a putea consulta pacienţii incompatibil. Iată încă un exemplu de neajutorare învăţată, dar cu consecinţe dintre cele mai grave.

Inteligenţa artificială a afectat şi alte domenii în mod bizar, în 2002, cercetătorii de la centrul Open Infrastructura for Network Computing al universităţii Berkeley şi-au dat seama că dacă am putea permite ca

calculatoarele noastre să lucreze pentru alţii în timp ce dormim, s-ar putea simula puterea unui super-calculator, iar acea putere ar putea fi folosită în scopuri ştiinţifice. Primele experimente s-au dovedit utile, iar sute de mii de oameni au donat timpul mort de procesare unor proiecte nobile din lume, cum ar fi Quake-Catcher Network, o reţea care caută activitatea seismică, sau SETI – home, care caută viaţa extraterestră în univers. În 2018, unii antreprenori inteligenţi şi-au dat seama cum aceste reţele ar putea fi reconfigurate pentru varianta actualizată a unei economii bazate pe sarcini de scurtă durată. În loc să conducă pentru Über sau Lyft, persoanele pot instala aplicaţii pe cont propriu prin care să câştige bani din timpul lor liber. Cele mai recente aplicaţii permit companiilor să folosească aparatele noastre, în schimbul unor credite sau sume de bani pe care le putem cheltui în alte scopuri.

La fel ca la începutul serviciilor de ride-sharing, mulţi oameni au părăsit forţa de muncă tradiţională pentru a lua parte la această variantă nouă a economiei bazată pe sarcini de scurtă durată. Şi-au dat demisia, şi au încercat să câştige bani închinând accesul la aparatele lor. Acest fenomen a provocat o presiune semnificativă asupra furnizorilor de curent electric şi internet, care nu au mai putut face faţă cererii. În zilele noastre, supraîncărcarea reţelei şi penele de curent sunt des întâlnite, iar din cauza faptului că oamenii au tendinţa de a rula aceste programe în timpul somnului, ei nu-şi dau seama că au pierdut un venit potenţial până dimineaţa.

Cei care au rămas în forţa de muncă tradiţională au început să folosească inteligenţa artificială pentru a-şi actualiza CV-urile şi scrisorile de intenţie, provocând o altă problemă. Problemele care ar fi putut elimina anumiţi candidaţi sunt acum mai puţin evidentetoţi par acum că au un avantaj competitiv. Sistemele de inteligenţă artificială sunt folosite pentru a tria candidaţii, dar managerii nu mai pot alege pentru că toţi candidaţii par la fel de buni.

Ei recurg aşadar la un comportament care să le dea o senzaţie de confort: bărbaţii albi au tendinţa să angajeze bărbaţi albi, pentru că sunt paralizaţi de tirania alegerii.

În majoritatea companiilor mari, ierarhia anterioară a fost redusă la două niveluri de muncitori: angajaţi cu aptitudini şi management superior. Angajaţii cu aptitudini lucrează împreună cu sisteme de inteligenţă artificială şi raportează unor superiori din inteligenţa artificială, pentru că întreaga pătură a managementului de mijloc a fost eliminată. La muncă, sistemele de supervizare monitorizează productivitatea, vă urmăresc în birou, notează cu cine socializaţi şi înregistrează nivelul vostru de fericire, anxietate, stres şi mulţumire. Sistemele sunt o întruchipare a acelor postere motivaţionale oribile, care vă aduc aminte că „sunteţi mai curajoşi decât credeţi” şi „sunteţi mai puternici decât pretextele”.

Guvernele au fost luate prin surprindere de eliminarea la scară largă a managementului de mijloc în industriile bazate pe cunoştinţe – cum ar fi dreptul sau finanţele pentru că s-au concentrat exclusiv pe munca sau ocupaţiile cu aptitudini reduse cum ar fi şoferii, fermierii sau muncitorii. Domeniile creative sunt lovite la fel de greu de o nouă ramură a inteligenţei artificiale: creativitatea maşinilor. Graficienii, arhitecţii, copywriterii şi dezvoltatorii web au devenit inutili pentru că reţelele antagonice generative şi sistemele de inteligenţă artificiale mai recente s-au dovedit a fi remarcabil de productive şi sigure. În acelaşi timp, inteligenţa artificială a conferit super-puteri anumitor posturidirectorii generali, financiari sau tehnici. S-a căscat o prăpastie semnificativă, şi vârful organizaţiilor concentrează din ce în ce mai multe averi. Constatăm începutul unui sistem de caste în epoca digitală.

Încă un defect: contaminarea informaţiilor. În urmă cu un deceniu, multe procese şi reglementări internaţionale au provocat fracturarea internetului. În locul unei singure reţele mondiale, am ajuns să avem mai

multe reţele paralele, în care regulile în domeniul digital depind de legile naţionale şi de restricţiile geografice. Procesul nu s-a produs imediat. Când internetul a trecut de la guvern şi mediul academic înspre mediu privat, în anii ’90, i-am permis să se răspândească liber, în loc să impunem reglementări, precum în domeniul utilităţilor sau cel financiar. La vremea aceea, legiuitorii nu s-au gândit prea mult la felul în care ar putea fi folosite ulterior toate datele pe care le generăm pe internet. Acum a devenit imposibil de respectat fiecare lege diferită a fiecărei ţări, iar filtrele anterioare de conţinut s-au extins pentru a corespunde graniţelor geografice. Astfel, s-a contribuit la promovarea şi propagarea ştirilor false. Deoarece actorii malefici folosesc algoritmii de generare şi pentru că primim ştiri diferite în funcţie de regiune, nu ştim în cine sau în ce să avem încredere. Toate marile surse de ştiri din lume au fost păcălite cel puţin o dată, deoarece până şi jurnaliştii cu experienţă întâmpină dificultăţi în a verifica atât videoclipuri cu lideri globali dar şi cu oameni obişnuiţi. E aproape imposibil să ne dăm seama dacă vocile şi chipurile din clipurile pe care le urmărim sunt simulate sau adevărate.

O altă problemă: apariţia unui val de criminalitate bazată pe inteligenţa artificială pe care nimeni nu l-a anticipat. Programe cu aplicaţii reduse dar puternice au început să provoace probleme pe internet. Acestea se ocupă cu achiziţii ilegale: genţi de marcă contrafăcute, droguri sau medicamente produse din animale braconate (cum ar fi coarnele de rinocer sau fildeşii de elefant). Aceste programe ne spionează pe reţelele de socializare, citesc ştirile şi infiltrează pieţele financiare provocând reprize bruşte de panică, în spaţiul public, aceste programe calomniază, cu intenţia de a defăima caracterul şi reputaţia oamenilor. Începem să ne facem probleme că inteligenţa artificială ne-ar putea accesa DDP-urile, modifica datele biometrice şi falsifica atât propriile date, cât şi cele pe care le-am moştenit. O parte a acestor

nelegiuiri sunt provocate intenţionat de către varianta modernă a mafiei: o reţea generalizată de criminalitate organizată dificil de depistat şi ţinut sub control. Unele dintre inteligenţele artificiale malefice sunt însă rodul întâmplării: ele au evoluat pur şi simplu şi au început să se comporte de o manieră pe care nimeni nu o plănuise.

Problemele se extind şi la roboţii fizici. Roboţii de securitate, dotaţi cu camere inteligente şi programe analitice predictive, au tendinţa să urmărească în mod regulat oamenii de culoare. Roboţii de securitate nu sunt dotaţi cu arme, dar emit comenzi cu voce foarte tare şi alarme cu sunete ascuţite dacă bănuiesc că se petrece vreo nelegiuire, în interiorul clădirilor de birouri, al hotelurilor, gărilor şi aeroporturilor, oamenii de culoare sunt hărţuiţi şi umiliţi în mod constant pentru că un robot de securitate i-a etichetat ca fiind suspecţi.

Mafia G nu are o relaţie uşoară cu forţele de ordine americane, care vor să acceseze DDP-urile. În loc să colaboreze, autorităţile ameninţă cu procese şi încearcă să oblige Mafia G să partajeze datele, iar companiile nu au nicio obligaţie – de ordin juridic sau altminteride a ceda acestor solicitări. Deşi nimeni nu vrea să declare acest lucru în public, se pare că autorităţile americane speră să poată replica anumite părţi ale monitorizării algoritmice şi a sistemului de credite sociale din China. Temându-se de represaliile consumatorilor, Mafia G continuă să refuze accesul la sistemele lor.

Am vorbit de peste un deceniu despre impactul filosofic şi etic al procesului decizional în cadrul forţelor de ordine, însă fără a se ajunge la vreun standard, vreo normă sau vreo reglementare. Avem acum un şir aparent infinit de infracţiuni comise cu ajutorul inteligenţei artificiale, fără a avea însă metode de pedepsire. Nu există închisori pentru inteligenţe artificiale sau roboţi. Legile care definesc infracţiunile nu se aplică şi tehnologiilor pe care le-am creat.

Confuzia şi dezamăgirea noastră se pliază de minune

pe strategia Chinei, care nu e doar un rival care tinde spre egalitatea cu Statele Unite, ci un concurent direct şi o ameninţare de ordin militar. China a petrecut decenii furând designul echipamentelor americane şi strategiile de apărare, iar această tactică a adus beneficii. Preşedintele Xi consolidează şi mai mult puterea armatei chineze, concentrată mai mult pe programare decât pe luptă fizică. Spre exemplu, spectacolele luminoase chinezeşti prezentate cu ocazia diferitelor evenimente – un festival din 2017 cu „lampioane cu drone”, o extravaganţă de vară din 2018 cu artificii cu drones-au dovedit a fi teste pentru supravegherea maselor. Armata Chinei foloseşte acum drone puternice bazate pe inteligenţa artificială care vânează în haite, pe uscat şi pe mare.

Prin intermediul puterii economice, diplomaţiei directe şi puterii militare, China a deschis o nouă epocă colonială, colonizând cu succes Zambia, Tanzania, Republica Democrată Congo, Kenya, Somalia, Etiopia, Eritreea şi Sudanul. China construieşte infrastructura-şi implementează sistemul de credite socialeşi extrage resurse naturale esenţiale pentru a-şi sugruma concurenţa şi a-şi alimenta clasa mijlocie aflată în expansiune. China controlează acum 75% din rezervele mondiale de litiu, o materie primă esenţială pentru baterii. China a distrus şi pădurile de santal şi a condus la dispariţia copacului Mukula, o specie cu creştere lentă din Africa centrală care fusese folosită o perioadă pentru a fabrica mese şi scaune sculptate colorate în roşu.

Nicio putere economicănici Statele Unite, Japonia, Coreea de Sud sau Uniunea Europeanănu are suficientă influenţă politică sau economică pentru a împiedica China să-şi extindă zonele de influenţă economică specială în Marea Chinei de Sud, Marea Chinei de Est şi Marea Galbenă. Aproape jumătate din comerţul global trebuie să tranziteze una din aceste zone, şi fiecare navă care tranzitează trebuie să plătească o taxă considerabilă guvernului chinez.

Unii analişti spun că Beijingul a ratat ţinta de a deveni principala putere în domeniul inteligenţei artificiale în 2025, chiar dacă a preluat controlul asupra anumitor resurse din lumea fizică. Aceşti analişti nu văd însă imaginea de ansamblu. După ani de contracte de transferuri obligatorii de tehnologie, practici de restrângere a concurenţei, şi investiţii considerabile chineze în companiile europene şi americane, strategia Chinei s-a dovedit extrem de reuşită. China domină în domeniul tehnologiilor avansate, inclusiv în ce priveşte robotica, sursele noi de energie, genetica şi aviaţiaiar toate aceste domenii se folosesc de inteligenţa artificială şi sunt îmbunătăţite cu ajutorul acesteia. Nu există date oficiale, dar luând în calcul laboratoarele de cercetare de stat, parteneriatele cu Baidu, Alibaba, şi Tencent şi partenerii din proiectul Noul Drum al Mătăsii, experţii estimează că China a reuşit să sporească valoarea ecosistemului inteligenţei artificiale cu peste 500 de miliarde de yuan (aproximativ 73 de miliarde de dolari) doar într-un deceniu.

2049: Şi au rămas doar cinci

Pe măsură ce timpul trece şi se înregistrează progrese în domeniul inteligenţei artificiale generale, constelaţia celor Nouă Titani s-a modificat de o manieră atât profundă cât şi problematică. Grupul BAT din China e acum mai puternic ca niciodată, şi continuă să primească ordine de la Beijing. Din cei şase membri originali ai Mafiei G, au rămas acum doar cinci, din cauza parteneriatelor strategice şi alianţelor: Amazon-Apple şi Google IBM sunt companiile cu cea mai mare putere. Microsoft oferă în prezent doar asistenţă tehnică pentru sistem şi servicii depăşite.

Poate cea mai surprinzătoare evoluţie este însă cea a Facebook. Nici scandalul Cambridge Analytica şi nici măcar dezvăluirile despre amestecul Rusiei în alegerile din Statele Unite nu au condus la prăbuşirea Facebook. Nu a fost nici oboseala pe care o resimţeam cu toţii, în timp ce

fluxurile se umpleau din ce în ce mai mult cu ură, teorii ale conspiraţiei politice şi demagogie. Modelul de afaceri al Facebook s-a dovedit a nu fi sustenabil pe termen lung. Odată ce utilizatorii au început să nu mai folosească reţeaua şi companiile de publicitate nu au mai cheltuit bani pe platformă, Facebook s-a găsit în postura de a nu avea venituri din mai multe surse. Până în 2035, compania a început să aibă probleme financiare grave. Acţionarii voiau să vândă, directorii fondurilor mutuale s-au speriat, şi piaţa s-a întors împotriva companiei, care a fost vândută pe bucăţi. Toţi cei ale căror date erau blocate în interiorul reţeleimajoritatea americanilor sunt acum foarte îngrijoraţi, pentru că datele noastre au fost cumpărate pe ascuns de un consorţiu. Ancheta e în desfăşurare, dar se pare că acest consorţiu e de fapt o companie controlată de China. E probabil ca acum să facem cu toţii parte din sistemul chinezesc de credite sociale şi că suntem cu toţii urmăriţi.

Şi voi, la fel ca ceilalţi americani, aţi învăţat să trăiţi cu o anxietate constantă, de nivel scăzut. Nivelul naţional de nelinişte este adesea comparat cu ameninţarea războiului nuclear din anii 1960 şi 1980. De data aceasta însă, obiectul concret al fricii este necunoscut. Nu ştim dacă DDP-ul este în siguranţă, sau la ce date are acces China. Sunteţi nesiguri cât de infiltraţi sunt hackerii chinezi în infrastructuratării noastre. Vă treziţi adesea dimineaţa devreme, întrebându-vă ce ştie China despre voi, despre podurile pe care mergeţi spre muncă, despre ţevile de gaz care vă intră în casă, şi ce plănuiesc să facă cu toate aceste informaţii.

Ceea ce nu am putut anticipa a fost un spectru de LAG, construite pentru diferite scopuri şi sarcini, care sunt atât puternice dar şi indiferente la valorile umane. Dacă privim în urmă, am fost extrem de naivi. Pe măsură ce Amazon, Apple, Google şi IBM au încheiat parteneriate, au ales o tabără sau alta şi au crescut, aceste companii nu şi-au setat şi standarde globale. Cu decenii în urmă, oamenii

cumpărau aplicaţii sau jocuri din magazinul Google Play, şi pentru că aproape oricui îi era la îndemână să lanseze şi să vândă o aplicaţie, calitatea diferea semnificativ. Existau mult prea multe aplicaţii care consumau bateria rapid, jocuri care colectau şi partajau informaţii personale şi reclame enervante care stricau experienţa pe mobil. Aceleaşi fenomene se manifestă acum în cazul IAG, însă cu consecinţe mult mai grave. Unele IAG se prefac că urmează protocoalele pe care le-au primit, apoi le suprascriu cu noi instrucţiuni. Unele IAG se îmbunătăţesc singure, chiar dacă cei care le-au creat nu le-au programat să facă asta. Unele se înmulţesc singure, se subdivid în alte IAG şi obţin resursele de care au nevoie pentru a-şi îndeplini obiectivul, indiferent de impactul pe care l-ar putea avea aceste acţiuni asupra ecosistemului.

Pentru a contracara problema LAG care nu se comportă cum trebuie, cercetătorii Applezon şi Google – IBM au implementat programe IAG dădacă – LAGD pe scurptentru a ţine în frâu alte sisteme. LAGD au un set clar de protocoale:

să investigheze şi analizeze alte LAG, pentru a vedea dacă acestea încalcă obiectivele stabilite iniţial să înregistreze în amănunt toate LAG care nu se comportă cum trebuie, dar şi întregul lor istoric (de ex., cine le-a creat, când au fost modificate, şi de către cine/ce)

să găsească primul programator şi să-l înştiinţeze de faptul că LAG nu se comportă cum trebuie după o perioadă de graţie (care depinde de gravitatea infracţiunilor IAG) să scoată din uz orice LAG să nu-şi modifice niciodată propriile obiective

Este evident că atât Applezon cât şi Google – IBM încercau să controleze un sistem care începea să scape de sub control, însă în afara ecosistemelor Applezon şi Google IBM, IAGD nu au fost adoptate la scară largă. Folosind ca precedent hotărârile anti-trust împotriva Google şi Microsoft, Parlamentul European a susţinut că LAGD nu sunt nimic mai mult decât o tentativă disimulată a

companiilor de a înăbuşi antreprenorii şi de a zdrobi concurenţa. UE a devenit prima care a interzis IAGD. Chiar şi în timp ce cercetătorii le cereau autorităţilor să permită acestor LAG specializate să contribuie la limitarea unei probleme serioase şi în curs de agravare, Congresul Statelor Unite a votat împotriva giganţilor din domeniul tehnologiei, interzicând folosirea LAGD. Aceste decizii lipsite de viziune nu au făcut decât să sădească o neîncredere a publicului în Applezon şi Google – IBM, care ar fi putut fi altminteri nişte păzitori de încredere ai DDP-urilor noastre.

Casa voastră a devenit un mare receptacul pentru marketing, care are un caracter constant şi invaziv. Vedeţi reclame personalizate oriunde există un ecran: în oglinzile inteligente din baie şi dulap, în ecranele retractabile din buzunare, şi chiar în geamurile inteligente pe care a trebuit să le instalaţi în case pentru a bloca radiaţiile solare extreme. Vă simţiţi incomod în propria casălocul unde odinioară eraţi cel mai relaxaţi.

Această neîncredere a transformat sistemul de sănătate într-un monstru. Applezonn Health Sistem şi Watson Calico au înregistrat progrese deosebite atât în domeniul inteligenţei artificiale cât şi al medicinei. Ambele au preluat ideea de la un costum robotic controlat cu mintea prezentat la Cupa Mondială din 2014. Neurobiologul Miguel Nicolelis de la Duke University şi-a dat seama cum să îmbine mintea şi maşina-iar studiile lui i-au inspirat şi pe alţii să scoată pe piaţă interfeţe creier-maşină. În anumite birouri unde tehnologia este mai avansată, muncitorii sunt încurajaţi să poarte bentiţe electronice şi să-şi conecteze minţile, împreună cu o IAG, pentru a rezolva probleme dificile. Nu toţi oamenii se simt confortabile cu această variantă hipertehnologizată a inteligenţei colective, deoarece interacţiunea presupune ca datele să treacă fie prin Applezon sau prin Watson Calico, care pot acum vedea la propriu ce ne trece prin cap.

Watson Calico a reuşit, în parteneriat cu o

universitate de top din New York, să ducă mai departe una dintre teoriile mai puţin cunoscute ale lui Turing legată de morfogeneză şi inteligenţa artificială. Turing credea că un sistem de substanţe chimice reacţionează probabil unele cu celelalte, iar acea reacţie se transmite de la o celulă la alta, modificând o parte dintre aceste celule. S-a dovedit că Turing a avut dreptate. S-au folosit sistemele IAG pentru a descoperi diferite moduri prin care să se creeze fiinţe multicelulare, dând astfel naştere unor fiinţe umane modificate, pe care le denumim „himere om-animal”.

Intenţia originală era de a crea ţesuturi omeneşti care să poată fi transplantate, aşa că am folosit porcii şi oile pentru a creşte ficaţi, inimi şi rinichi. Cercetătorii au dezvoltat şi ţesut neuronal. Domeniul părea promiţător, până când ne-am dat seama că IAG era folosită pentru a dezvolta himere om-animal care aveau alte caracteristici cum ar fi porci implantaţi cu ţesut cerebral omenesc care ajungeau la un IQ similar cu al unui om neinteligent şi copii nou-născuţi cu un simţ al mirosului ca de câine. Ceea ce nimeni nu a dezbătut (sau stabilit) sunt implicaţiile pe care le presupun aceste atribute himerice, care sunt moştenibile. Ce se întâmplă atunci când un om modificat pentru a avea capacităţi extrasenzoriale are un copil cu o persoană care are la rândul ei modificări?

Cel mai îngrijorător este faptul că China a decis să le dea un nou scop interfeţelor creiermaşină şi IAG ele fiind menite iniţial să ajute bolnavii să-şi recapete facultăţile în scopul obţinerii unor avantaje strategice de ordin militar. Sistemele au fost folosite pentru a spori capacităţile cognitive ale soldaţilor, care luptă în mare măsură din interiorul unor buncăre subterane întunecoase. În Statele Unite şi Uniunea Europeană, aceste experimente şi aplicaţii ale tehnologiei încalcă legile privind etica.

Din cauza practicilor coloniale chineze, a expansiunii zonelor de influenţă economică şi a folosirii lipsite de scrupule a IAG, începem să constatăm un declin real al civilizaţiei occidentale şi al idealurilor noastre

democratice. Sănătatea economiei noastre este în pericol, iar indicatorii tradiţionali precum numărul de construcţii noi sau vânzările de alimente şi ale magazinelor scad de la un trimestru la altul.

Chiar şi Applezon şi Google – IBM înregistrează într-un final o scădere a veniturilor, companiile fiind îngrijorate în privinţa viitorului. În timp ce lucrează pentru a moderniza DDP-urile pentru ca ele să poată lucra cu IAG paznici, ambele companii detectează o serie de date ciudate în sistemele de înregistrare. Există fragmente de cod care par să nu aibă sens, iar o parte din IAG care procesează şi direcţionează DDP-urile noastre se mişcă cu o viteză redusă. Într-un gest rar de colaborare, cele două companii partajează rezultatele analizelor în speranţa de a putea stabili problema. La birou şi acasă, luminile se închid la întâmplare. Ochelarii noştri inteligenţi au perioade intermitente în care nu funcţionează. Sateliţii noştri de comunicaţii ies de pe orbită.

Deşi nu le putem auzi, ştim că primele lovituri de tun au fost trase, şi că China a declarat război Americii.

2069: Ocupaţia digitală a Statelor Unite ale Americii

Ne dăm seama că China şi-a dezvoltat o generaţie de LAG care au capacităţi mult mai mari decât până în prezent. În absenţa unor LAGD care să păstorească LAG care o iau razna, China a putut construi şi lansa un sistem îngrozitor prin care să controleze majoritatea oamenilor de pe Pământ. Dacă nu acceptăm pretenţiile Chinei, aceasta ne va deconecta sistemele de comunicaţii. Dacă nu oferim Partidului Comunist Chinez accesul la datele noastre, acesta va închide infrastructura de bază, precum centralele electrice sau controlul traficului aerian.

Sunteţi un cetăţean în Statele Unite ale Americii sub ocupaţie digitală chineză. Transportul, banca, sistemul de sănătate, comutatoarele de lumini şi frigiderele sunt toate la cheremul Chinei.

Ceea ce a început ca o incursiune colonială în Africa a generat apariţia unui nou imperiu chinez, construit pe

baza inteligenţei artificiale. Omenirea se află în pragul unei SLA dezvoltate de către oţ ară care nu împărtăşește valorile şi idealurile noastre democratice.

DINASTIA RENGONG ZHINENG: SCENARIUL CATASTROFAL

„Aşa se sfârşește lumea, nu cu mare zgomot, ci cu un oftat”. – T.S. ELIOT

Până în 2023, am ignorat complet parcursul evolutiv al inteligenţei artificiale. Am ratat toate semnalele de alarmă, le-am ignorat şi nu ne-am imaginat în mod activ viitorul. I-am ajutat pe cei nouă Titani să concureze împotriva lor înşiși, ne-am dat frâu liber dorinţelor de consum, cumpărând cele mai recente aparate, sărbătorind orice nouă ocazie de a ne înregistra vocile şi chipurile, şi transmiţând aceste informaţii printr-o reţea care neaabsorbit în mod continuu datele. Am partajat videoclipuri stupide cu Alexa care nu poate avea o conversaţie corectă cu copiii noştri. Am cerut ca televizoarele să poată să ne scaneze feţele, fără să ne punem problema ce nevoie ar putea avea un televizor de datele noastre biometrice. De fiecare dată când Google lansa câte un nou proiect amuzant care ne mapa corpul în fotografii, chipul în tablouri, vocea pe a unor celebrităţi, amprentele noastre pe oameni din ţări îndepărtate, şi irisul pe strămoşii noştri, am luat parte cu interes, disperaţi să ţinem pasul cu influencerii digitali şi cele mai recente meme.

Triburile inteligenţei artificiale spun că diversitatea contează. Aceasta e mantra lor. Ei repetă în mod constant această afirmaţie, la lansări şi conferinţe, la interviuri de angajare şi la şedinţele consiliilor de administraţie, pe Twitter sau în articole. O spun în pliantele facultăţilor. O spun prin postere atrăgătoare atârnate în lifturi sau lipite pe holurile birourilor. Tribul inteligenţei artificiale, în mare parte format din bărbaţi albi, este antrenat să recite această mantră în sălile de curs, în laboratoare şi în spaţiile de lucru. În loc să facă alegeri dificile şi schimbări,

aceştia repetă mantra şi promit că schimbarea va veni curând. Iar mantra funcţionează chiar în scopul în care a fost inventată: să elimine negativitatea din minte şi să facă triburile inteligenţei artificiale să se simtă mai bine în propria piele. Diverşi guru ai triburilor inteligenţei artificiale transmit mantra fiecărei noi grupe de discipoli, care trăiesc senzaţia unei mari reuşite atunci când o repetă.

Mantra este repetată în bula comodă a triburilor inteligenţei artificiale, care cred că ei promovează diversitatea în timp ce fac exact opusul. Ei promovează diversitatea de toate felurile – politică, religioasă, sexuală, de gen, rasială, etnică, de statut economic şi vârstăânsă nu fac niciun demers serios pentru a fi mai inclusivi. În loc să vedem un spectru mai colorat de oameni şi păreri în domeniul inteligenţei artificiale prin poziţii de vârf în mediul universitar, cercetare sau în Mafia G, nu vedem nicio schimbare.

Pe măsură ce perspectivele triburilor asupra lumii devin din ce în ce mai înguste, problemele pe care le-am amintit deja se înmulţesc. Accidentele şi greşelile continuă să devină mai frecvente, de pildă sistemele de vizualizare digitală care identifică în mod eronat persoanele de culoare şi le acuză de comiterea unor infracţiuni. Supravegherea se extinde devenind în acelaşi timp mai puţin evidentă. Linia de demarcaţie dintre datele personale şi cele pe care le generăm la muncă se estompează, la fel şi criteriile pe baza cărora se decide cine foloseşte datele noastre, şi când. Transparenţa sistemelor de inteligenţă artificială dispare cu totul. (Nu că ar fi fost vreodată grozavă.)

Mafia G sunt singurii proprietari ai datelor voastre personale, care ajung să includă toate aspectele vieţii: conţinutul e-mailurilor, mesajele pe care le trimiteţi copiilor, furniturile digitale pe care le lăsaţi în urmă căutând scaunul perfect de birou, contururile unice ale degetelor şi feţei voastre, unde vă plimbaţi şi cât de

repede mergeţi, peste cine daţi la magazin, dacă aveţi gripă şi ce medicamente luaţi. Algoritmii iau decizii în numele vostru folosind toate datele voastre. Ei stabilesc dacă primiţi sau nu o reducere când rezervaţi un zbor. Ei vă ajută sau vă împiedică să obţineţi o slujbă, vă permit să vă cumpăraţi o casă sau o maşină, vă oferă parteneri pentru prima întâlnire şi le comunică doctorilor dacă minţiţi sau nu în privinţa consumului de alcool, tutun sau a exerciţiilor fizice. Deoarece Google, Amazon, Apipe, Facebook, Microsoft şi IBM deţin aceste dateşi pentru că iubim produsele lor, chiar dacă nu avem încredere deplină în aceste companiinu vedem că faptul că aceste companii deţin controlul complet asupra DDP-urilor noastre reprezintă doar o variantă americană a sistemului de credite sociale din China.

Ne aflăm blocaţi într-un sistem de caste digitale, în care inteligenţa artificială face alegeri şi judecăţi bazate nu doar pe felul în care noi ne trăim viaţa, ci şi pe DDP-urile părinţilor şi rudelor noastre. Averea nu mai contează. Statutul este stabilit de un „comportament exemplar” iar definiţia „exemplarului” a fost dată acum mult timp de o minoritate de programatori care credeau că cursurile de yoga de la prânz, vizitele dese la chiropracticieni şi o dietă organică ketogenică. Dacă nu faceţi săptămânal o saună cu infraroşii, sistemul de inteligenţă artificială la care sunteţi legat va înregistra abaterea în DDP. Acest act de răzvrătire nu vă va afecta doar pe voi, pentru că dosarul personal este conectat la toate persoanele pe care le cunoaşteţi şi cu care sunteţi înrudite. Nu puteţi scăpa de păcatele celor din jur.

În viitorul apropiat, Amazon şi IBM vor încerca să convingă guvernele Statelor Unite, Regatului Unit, Germania şi Japonia să ofere accesul la o serie de date de sănătate ale cetăţenilor. Apple, Google, Microsoft şi Facebook vor întâmpina probleme şi mai mari din cauza proceselor anti-trust anterioare. Experimentele Amazon şi IBM se vor dovedi utile agenţiilor guvernamentale, care

vor oferi mai multe contracte profitabile pentru toată Mafia G.

În 2008, când multe ţări ale lumii au intrat într-o criză financiară provocată de o bulă speculativă imobiliară, China s-a arătat dornică să cumpere fier, petrol şi cupru din America Latină, protejând ţările din regiune de daune economice mai mari. Când preţul petrolului a scăzut în 2011, China a fost dispusă să investească în America Latină şi să salveze regiunea de la faliment. 179 în 2013, China a efectuat exerciţii militare comune pe coasta Braziliei-repetând manevra în 2014 pe coasta Chileului.180 în 2015, ministerul apărării din China a găzduit un summit de 10 zile dedicat logisticii militare, cu invitaţi din 11 ţări latino-americane, invitând apoi în anii următori ofiţeri din aceste ţări să participe la programele de dezvoltare a carierei în China.181 în timp ce guvernul american bate în retragere şi joacă un rol din ce în ce mai mic pe scena lumii, China pare axată pe expansiune, încheind înţelegeri cu ţări din Asia de sud-est, Africa – şi, desigur, din America Latină.

După un deceniu în care şi-a construit temeinic relaţiile în America Latină, China – şi nu Statele Uniteeste cea care furnizează echipament militar Venezuelei, Boliviei, Perú-ului şi Argentinei, inclusiv avioane şi arme.182 China are un motiv pentru care să înfiinţeze baze în tot hinterlandul Statelor Unite. În Patagonia, China a construit o antenă militară şi o staţie de control spaţial, precum şi un centru de urmărire a sateliţilor în nord-vestul Argentinei.183 Toate aceste activităţi presupun folosirea inteligenţei artificiale.

Politicienii nu sunt capabili să facă legătura între China, Statele Unite şi inteligenţa artificială. Consolidarea puterii lui Xi Jinping, diferitele proiecte ale statului chinez, economia aflată în creştere galopantă şi succesul grupului BAT reprezintă o forţă de neoprit – chiar dacă invizibilăşi care trebuie confruntată. Nici Casa Albă şi nici Congresul Statelor Unite nu văd că imixtiunile Chinei în toate acesteţ

ari-Tanzania, Vietnam, Argentina, Boliviaprivesc atât raţiuni economice *cât şi* de informaţii. Ei refuză să recunoască că China construieşte un imperiu al secolului 21 bazat pe date, infrastructura inteligenţei artificiale, geopolitică şi economia globală. E o eroare gravă de judecată pe care o vom regreta cu toţii ulterior.

Cetăţenii chinezi învaţă să trăiască cu monitorizarea automată şi consecinţele încălcării regulilor. Criminalitatea se reduce, şi neliniştea socială este înăbuşită, iar pentru o perioadă păturile superioare şi cele de mijloc îşi păstrează bunăstarea. Au acces la haine şi genţi de lux, mobilă de designer şi maşini elegante pe care părinţii sau bunicii lor nici nu şi le-ar fi putut imagina. Se fac promisiuni că într-o bună zi sărăcia va fi eliminată din China. Pentru moment, se pare că renunţarea la intimitate, libertate religioasă, identitate sexuală şi libertatea de expresie reprezintă un preţ rezonabil de plătit în schimbul unui scor bun al creditelor sociale.

Liderii politici americani nu se interesează suficient în privinţa a ce este, ce nu este şi de ce este importantă inteligenţa artificială. În afara discuţiilor obişnuite legate de perturbarea productivităţii şi a şomajului, conducătorii de la Washington nu fac niciun efort de a comunica serios cu Mafia G în privinţa celorlalte aspecte legate de inteligenţa artificială, cum ar fi siguranţa naţională, echilibrul geopolitic, riscurile şi oportunităţile inteligenţei artificiale generale, sau introducerea inteligenţei artificiale în alte domenii (cum ar fi genetica, agricultura şi educaţia).

În absenţa unei direcţii strategice referitoare la inteligenţa artificială din partea Casei Albe de fapt, în prezenţa unei atitudini ostile faţă de ştiinţă şi tehnologieautorităţile de la Washington se concentrează mai mult asupra aspectelor relevante prin prisma următorului ciclu electoral şi a veştilor care vor fi bine primite de showurile politice de duminică.

Nici Mafia G şi nici conducătorii acestor companii nu

iau măsuri care să pericliteze deliberat democraţia. Protejarea statului de superputere globală dominantă al Americii şi asigurarea conservării idealurilor democratice nu fac însă parte din valorile de bază ale acestor companii. La începutul anilor ’2010, fostul director al Google Eric Schmidt lucra intens şi admirabil pentru a spori gradul de pregătire al armatei şi guvernului Statelor Unite în epoca inteligenţei artificiale. Nu era vorba de o tentativă de a obţine contracte guvernamentale pentru Google. Schmidt era preocupat de siguranţa naţională şi de capacitatea de reacţie militară în această nouă eră tehnologică. Demersul era însă atât de neobişnuit încât companiile din Silicon Valley i-au pus în discuţie motivele. În loc ca ceilalţi conducători ai companiilor din Mafia G să-i urmeze exemplul, aceştia s-au arătat sceptici vizavi de ambiţiile lui. Prin urmare, în afară de Schmidt, niciun alt conducător al Mafiei G nu s-a mai gândit la rolul pe care îl joacă inteligenţa artificială în planul Chinei de a uzurpa poziţia dominantă a Americii.

Între Mafia G şi agenţiile guvernamentale sau militare nu există o colaborare strategică-cel puţin nu în afara unor contracte lucrative. Mafia G acceptă politicile complicate şi demodate de achiziţii ale armatei şi guvernului, însă acest fapt nu contribuie la o evoluţie a inteligenţei artificiale în interesul naţional. Această acceptare nu face de fapt decât să releve diferenţele culturale dintre Silicon Valley şi Washington, încetinind ritmul modernizării. Puţinele agenţii guvernamentale care au ca obiect încurajarea inovaţiei-united States Digital Service, Futures Command a Armatei Statelor Unite, comisia Defense Inovation Board şi iniţiativele Defense Inovation Unit Experimental (DIUx) – sunt fragile şi supuse riscului de reducere a personalului şi finanţării, în funcţie de cei care se perindă la putere. Washingtonul consideră că are o relaţie tranzacţională cu Mafia G. Nici parlamentarii şi nici Casa Albă nu încearcă cu adevărat să dezvolte relaţiile necesare cu conducerea Mafiei G pentru

a forma o coaliţie pe termen lung în domeniul inteligenţei artificiale. Mafia G, armata şi guvernul Statelor Unite îşi dau târcoale fără să-şi unească eforturile în interesul naţional.

Am dat voie ca egoul şi obiceiurile să împiedice degajarea unui consens naţional în privinţa Chinei. Politicienii, reprezentaţii comerciali, jurnaliştii, informaticienii şi profesorii au dezbătut problema relaţiilor dintre China, Statele Unite şi inteligenţa artificială în continuu, fiecare parte susţinându-şi convingerile preferate, tradiţionale, fără să lase loc unor realităţi alternative. Suspecţii de serviciu susţin că Xi nu va rămâne prea mult timp la putere, chiar şi dacă s-au abrogat limitele privind numărul de mandate. Odată ce Xi va pierde puterea, toate planurile pe termen lung ale Chinei se vor nărui. Alţii combat această perspectivă: Xi va uni poporul şi partidul. Indiferent dacă moare înainte de vreme sau îşi cedează poziţia unui succesor, Partidul Comunist Chinez va fi mai puternic şi va duce la îndeplinirea planurile privind inteligenţa artificială. Discuţia se poartă tot aşa, înainte şi înapoi: politicile industriale chineze nu vor avea niciun impact-sau vor provoca distrugerea economiei Statelor Unite. Armata Chinei ameninţă lumea occidentală-sau e doar un articol exagerat de care ne vom plictisi curând. Ar trebui să investim timp şi bani într-o strategie naţională privind inteligenţa artificială, ştiind că planurile Chinei ar putea de greş-sau mai bine ne-am păstra timpul şi banii şi am adopta o abordare bazată pe aşteptare. Un singur aspect este acceptat de toată lumea: dacă America are cu adevărat probleme, Mafia G va fi obligată să ne ajute.

Politicienii şi grupurile noastre de gândire susţin aceleaşi argumente obosite, fără să acţioneze, menţinându-se într-o stare de inacţiune. Aleg această variantă pentru că în America e greu să scapi de forţa centripetă a profitului fără o intervenţie puternică.

Am auzit povestea acestei inacţiuni de multe ori până acum. Am apărat statutul fumatului, dezbătând datele

concrete despre cancer şi continuând să vindem tutunul ca un accesoriu al femeilor la modă, un stimulent pentru muncitori şi un remediu medicinal pentru oameni bolnavi. Nu am luat măsuri în privinţa încălzirii globale, certându-ne constant în jurul unor definiţii. Dacă încălzirea globală e reală, cum de poate fi atât de *rece*? Ne-am resemnat să dezbatem etapele. Afirmaţiile alarmante ale anilor ’70 au devenit sumbre în anii ’90 apoi apocaliptice în anii ’2010, dar suntem cu toţii în continuare aici. Cine zice că lucrurile vor deveni atât de rele în viitor?

Schimbarea sistemică are un efect de compunere şi se acumulează în decenii, nu în zile. Când ne vom da seama că inacţiunea a reprezentat alegerea greşită, va fi fost deja prea târziu.

2029: Prizonieri digitali în ultimul deceniu, aţi fost stimulaţi să cumpăraţi tot soiul de tehnologii şi sisteme bazate pe tehnologia artificială. Toate aparatele electrocasnice sunt dotate cu astfel de sisteme. Frigiderul ţine cont de mâncarea pe care o răceşte. Maşinile de spălat – chiar şi cele de la spălătoriiurmăresc ciclul de spălare şi vă anunţă când s-a terminat. Cuptorul se închide singur înainte ca curcanul să se usuce şi să ia foc.

Există însă şi o problemă pe care nu aţi anticipat-o: nu aveţi permisiunea de a bloca funcţionarea unui sistem de inteligenţă artificială care se voia a fi „de ajutor”. După ce puneţi prăjiturile, berea, brânza şi carnea în frigidertoate cumpărate pentru ziua meciului Super Bowlfapta este consemnată în DDP. Numărul de porţii şi calorii depăşeşte cel normal pentru numărul de oameni din casă, prin urmare inteligenţa artificială concluzionează că plănuiţi să mâncaţi peste măsură. Ar putea fi dincolo de miezul nopţii şi poate că aţi planificat deja să puneţi mâine rufele la spălat şi uscat înainte de a pleca la muncă, însă inteligenţa artificială a maşinii de spălat nu ia în calcul dorinţa voastră de a dormi. Ea sună o alarmă şi vă trimite constant notificări, fără să se oprească, pentru a vă spune că a venit vremea să puneţi hainele la uscat. Aţi vrea să

faceţi un mezel de curcan, dar cuptorul nu vă permite, pentru că inteligenţa artificială a fost programată să permită doar o carne suculentă. (Sau, dacă vă puteţi permite, puteţi plăti în plus pentru funcţionalitatea care să vă permită să faceţi şi mezeluri.)

Unele persoane au probleme legate de inteligenţa artificială, mai ales aparatura de bucătărie, şi de regulă dimineaţa. Panourile de control se sting, intermitent, ceea ce din nefericire are ca efect închiderea uşii, împiedicându-vă să puteţi mânca. Maşina de spălat vase se opreşte subit, la mijlocul ciclului, ţinând paharele şi argintăria să băltească într-o apă soioasă, plină de grăsimi. Volumul difuzorului inteligent devine dintr-odată foarte ridicat, făcând imposibilă conversaţia cu membrii familiei la micul dejun. Şi voi, şi zeci de mii de alţi consumatori reclamaţi aceste defecţiuni, iar Mafia G însărcinează de fiecare dată câţiva manageri de produs să studieze problema. Jurnaliştii atribuie aceste defecte „felului dubios” în care „inteligenţa artificială se comportă uneori ciudat”.

La început, atacurile par aleatorii, discrete. Vom da deci cu toţii vina pe Google, Apple şi Amazon pentru produsele defecte şi serviciile detestabile de relaţii cu clienţii. Experţii în securitate informatică sunt apoi uluiţi să descopere că toate aceste defecţiuni sunt de fapt interconectate. E un nou tip de atac cibernetic pornit din China, şi efectuat cu ajutorul învăţării maşinilor. Chinezii i-au dat şi un nume: e în, sau „bei kun”, care în traducere înseamnă „prizonier”. Hackerii, susţinuţi de guvernul chinez, au crezut că e o idee bună să ne atace la micul dejun şi să ne ţină mâncarea, băutura şi tacâmurile blocate în aparatele noastre bazate pe inteligenţa artificială. Ei au un scop unic, şi complicat: să sădească neîncrederea în Mafia G.

Microsoft şi IBM continuă să existe, însă joacă un rol minor în spaţiul inteligenţei artificiale. Microsoft, o companie care publica odinioară studii de mare valoare în

domeniul vizualizării digitale sau procesării limbajului natural, nu a reuşit niciodată să-şi canalizeze eforturile interne pentru a putea concura în domeniul inteligenţei artificiale. Acum, compania îşi reduce dimensiunile şi oferă în principal asistenţă pentru sistemele sale învechite: ceea ce a mai rămas din serviciile cloud Azure, Sharepoint, Skype şi Outlook. În timp ce IBM Watson a găsit parteneri şi clienţi, serviciul cloud al IBM, care ocupase mult timp de la mare distanţă locul al treilea după Amazon şi Microsoft, s-a micşorat odată ce Google a început să ofere preţuri competitive atât agenţiilor guvernamentale cât şi marilor companii. Celelalte divizii: centrele de date, stocare şi semiconductori – nu au putut lupta împotriva concurenţilor din Taiwan, deveniţi între timp cei mai mari furnizori din lume. Pentru companiile din Taiwan, principiul „O singură Chină” al Partidului Comunist Chinez se traduce printr-un avantaj semnificativ de piaţă, chiar dacă Beijingul le restricţionează libertăţile individuale. Politica industrială a Chinei a reuşit să împiedice IBM să vândă în multe zone de pe glob.

Ce s-a ales de Facebook? După ani în care a promis că-şi va spori securitatea şi va oferi o transparenţă mai mare asupra felului în care partajează datele noastre, majoritatea utilizatorilor au ales alte platforme. Copiii din generaţia Alpha (copiii generaţiei Y) au umplut reţeaua cu pozele lor, fără însă să-şi facă conturi. Facebook o ia încet pe calea dispariţiei, după My Space.

În 2035 interoperabilitatea reprezintă încă un punct slab în ecosistemul inteligenţei artificiale din Occident, utilizatorii fiind forţaţi să accepte un sistem bazat pe o segregare de facto. Aparatele noastre sunt conectate la Google, Apple sau Amazon, aşa că avem tendinţa de a cumpăra doar produsele şi serviciile oferite de una dintre aceste trei companiicompanii care ne-au vândut şi toate aparatele bazate pe inteligenţă artificială din casele noastreavem familii Google, familii Apple sau familii Amazon. Aceste etichete vin la pachet şi cu prejudecăţi

nedorite.

Gospodăriile Apple au tendinţa de a fi mai în vârstă, şi mai bogate. Acestea îşi pot permite toate aparatele Apple, disponibile într-una din cele trei culori: un alb-argintiu, gri, sau onix închis. Ochelarii, toaletele şi frigiderele Apple continuă tradiţia lungă a produselor scumpe pe care oricine le poate folosi imediat după ce le scoate din cutie. DDP-urile Apple sunt dotate cu interfeţe vorbitoare şi două voci diferite, Joost (cu o „tonalitate unisex mai ridicată”) sau Deva (cu o „tonalitate unisex mai gravă.). Uşurinţa de utilizare vine la pachet cu un cost. Inteligenţa artificială a Apple nu poate fi suprascrisă. Într-o casă Apple în care funcţionează aerul condiţionat, nu puteţi lăsa uşa deschisă mai mult de un minut fără ca sistemul să nu piuie fără încetare. Dacă senzorii din becuri detectează o cantitate suficientă de lumină de zi, sistemul Apple va împiedica acţionarea comutatorului de lumină.

Am văzut în avanpremieră sistemul Google pentru case conectate acum câteva decenii, la festivalul South By Southwest din 2018 din Austin, Texas. În vremea aceea, replica era „Make Google do it”, iar modelele atrăgătoare invitau grupuri mici de oameni într-o casă cu trei etaje în care puteau să interacţioneze cu aparate cu ecrane bazate pe inteligenţa artificială şi blendere conectate. Sistemul Google e mai puţin intuitiv, dar foloseşte DDP-urile noastre mai bine-oferind niveluri diferite de deservire şi acces. Pentru cei care-şi pot permite tarifele de acces la o categorie mai bună şi au cunoştinţe mai bune în domeniul tehnologiei, Google Green oferă familiei capacitatea de a-şi debloca manual sistemele, putând conecta mai multe aparate – cum ar fi cafetiere, imprimante 3D, sisteme de irigaţii din grădinăla propria casă. Familiile Green au opţiunea de a renunţa la marketing şi reclame, deşi datele lor sunt în continuare strânse şi trimise unor terţi. Google Blue reprezintă o variantă accesibilă în care drepturile de deblocare sunt limitate, însă familiile Blue primesc în continuare conţinut de marketing. Google Yellow

reprezintă nivelul cel mai de jos. Serviciul este gratuit, dar nu oferă capacitatea de schimbare a setărilor, oferă o selecţie mică de aparate şi o protecţie limitată a datelor.

Amazon a ales o direcţie interesantă dar care s-a dovedit în cele din urmă profitabilă. Câteva anunţuri făcute de Amazon în toamna lui 2018 au trecut în mare parte neobservate, cum ar fi lansarea cuptorului cu microunde Amazon Basics, care include o interfaţă vocală. Utilizatorii ar putea pune o pungă cu pop-corn în microunde şi să ceară Alexei să o prepare. Jurnaliştii din domeniul tehnologiei au considerat că acest cuptor cu microunde constituie un mod nou şi prostesc de a o folosi pe Alexa, ignorând imaginea de ansamblu: sistemul fusese de fapt creat pentru a ne face dependenţi de abonamentul la pop-corn. Acel cuptor monitorizează atât ce încălzim dar şi ce comandăm de pe platforma Amazon. O nouă cutie soseşte înainte ca cea veche să se termine.

Fiindcă Amazon a avut cea mai inteligentă abordare de colaborare cu agenţiile federale, statale şi localeoferindu-le reduceri semnificative pe Amazon.com, acceptând cu răbdare condiţiile de achiziţii, construind şi oferind mentenanţa pentru servicii de cloud –, compania a reuşit să devină platforma preferată pentru oferirea anumitor servicii sociale în Statele Unite. Amazon a descoperit astfel cum să extragă bani de la guvern.

Familiile cu venituri scăzute trăiesc în Amazon Housing, care a înlocuit programele de locuinţe ale autorităţilor din Statele Unite. După orice standard, locuinţele sunt mult mai bune decât cele oferite anterior prin programele guvernamentale. Casele Amazon sunt dotate în fiecare cameră cu aparate conectate. Fostul program naţional de asistenţă nutriţională (cunoscut sub numele de Programul Bonurilor de Masă) este oferit în prezent prin Amazon, care oferă produse electrocasnice marca Amazon la preţuri reduse semnificativ, precum şi alimente, cosmetice şi cărţi. Nu e de mirare că acest program funcţionează fără niciun fel de problemă. Nu

există niciodată întârzieri în distribuţia fondurilor, contul poate fi accesat cu uşurinţă şi toate tranzacţiile pot fi efectuate fără vreo coadă la un ghişeu. Cei care locuiesc în casele Amazon Homes trebuie să-şi cumpere majoritatea produselor prin Amazon, în timp ce datele pe care le produc sunt filtrate şi monetizate în diferite scopuri. Inteligenţa artificială a Amazon este cea mai răspândită, urmărind peste tot familiile Amazon pentru a colecta date comportamentale valoroase.

Imposibilitatea interconectării diferitelor structuri şi sisteme de inteligenţă artificială a condus la o segregare în funcţie de DDP şi locuinţă, motiv pentru care avem în prezent un sistem de caste digitale. Alegând Google, Apple sau Amazon, suntem forţaţi să aliniem valorile familiei cu cele ale corporaţiei. Familiile Apple sunt bogate, poate cu cunoştinţe mai modeste în domeniul inteligenţei artificiale şi trăiesc în case frumoase. Familiile Google ar putea fi bogate şi cu o înţelegere mai mare asupra tehnologiei, sau poate din clasa de mijloc şi cu o toleranţă mai mare faţă de marketing, sau acceptă faptul că nu e atât de important ca în viaţă să ai la dispoziţie multe alegeri. Despre familiile Amazon se poate spune pur şi simplu că sunt sărace, chiar dacă au acces gratuit la aparate cu puteri sporite.

Familiile sunt blocate în funcţie de DDP-uri, iar etichetarea le urmăreşte. E mai uşor pentru o familie Google Yellow să treacă la nivelul Blue sau Green decât pentru un membru dintr-o familie Amazon să se conecteze la sistemul Apple. Acesta este motivul pentru care majoritatea familiilor au ales Google, atunci când au avut ocazia. Statutul este vizibil pentru orice inteligenţă artificială cu care interacţionaţi. Taxiurile care se conduc singure, oferite de Lyft, über şi Citilar, nu preiau călători Amazon prea uşor, iar maşinile alocate lor au tendinţa de a nu fi la fel de bune. Maşinile Waymo deservesc doar clienţi Google. Pentru cei din categoria Green, maşina oferă temperatura dorită de călător şi iluminatul în culorile preferate, conducând de-a lungul traseelor preferate de

călător. Cei din categoria Yellow trebuie să îndure reclame pe toată durata călătoriei.

Reclamele nu sunt singura durere de cap pentru clienţii Google din categoria Yellow. Un dezavantaj al tuturor aparatelor subvenţionate (sau gratuite) oferite familiilor Google Blue, Yellow şi Amazon este că alertele inteligenţei artificiale legate de sănătate şi bunăstare nu pot fi deconectate, iar aceste sisteme monitorizează, diagnostichează şi sugestionează în mod constant. Când au fost iniţial construite, informaticienii au trebuit să definească din necesitate sănătatea şi bunăstarea pe baza unor criterii rigide. Acum, valorile triburilor inteligenţei artificiale din trecut sunt o amintire oprimantă dintr-o epocă mai simplă. Nerespectarea alertelor legate de sănătate şi bunăstare atrage multiple consecinţe.

Mai ţineţi minte sistemul Amazon Locker, acel dulap pe care le foloseaţi în urmă cu mulţi ani pentru a vă ridica cumpărăturile din aplicaţia Amazon şi de pe site-ul Amazon? El fac parte acum din casele Amazon. Departamentul de Sănătate şi Servicii Sociale al Statelor Unite a considerat că sugestionarea reprezenta o variantă inteligentă de a îmbunătăţi sănătatea şi bunăstarea oamenilor săraci, aşa că acest departament a elaborat noi politici prin care solicită tuturor ocupanţilor din locuinţele sociale să fie dotaţi cu tehnologia Locker. Aceste dulapuri ar putea părea ca nişte cămări, frigidere sau şifoniere obişnuite, însă ele funcţionează ca nişte jurii bazate pe inteligenţa artificială. Dacă un client Amazon Housing nu şi-a îndeplinit norma de exerciţii fizice, sistemul Locker va decide să-i încuie congelatorul şi nu-l va lăsa să mănânce îngheţată.

Simţim consecinţele negative ale lucrurilor care ne oferă plăcerea şi în afara caselor noastre Apple, Amazon şi Google. Bordelurile tehnologizate, în care lucrează roboţi sexuali bazaţi pe inteligenţa artificială, au căpătat o acceptare socială deoarece oferă o alternativă igienică la raporturile sexuale cu alţi oameni. Bordelurile operează pe

propriile platforme şi necesită un abonament care vă permite să construiţi şi educaţi o personalitate bazată pe inteligenţă artificială. (Sau chiar personalităţi, pentru cei care-şi permit pachetul premium). Alegeţi pur şi simplu un corp şi vă uitaţi în ochii lui camerele inteligente de dimensiuni reduse vă recunosc faţa. Odată ce robotul se trezeşte, el discută cu voi de parcă nu ar fi trecut timpul şi reacţionează la toate dorinţele şi ordinele voastre. Sexul obişnuit, cu persoane normale vă dezamăgeşte.

Nu e imposibil să vă căsătoriţi în afara castei din când în când o persoană Amazon se poate căsători cu un membru al unei familii Apple-însa dictonul cu „persoanele diferite se atrag” nu se mai aplică. Toate serviciile electronice matrimoniale ne aleg parteneri în funcţie de DDP şi de statut. Pe de o parte, nu mai suferim de tirania alegerii, deoarece sistemele de selecţie bazate pe inteligenţa artificială au redus dramatic numărul potenţialilor parteneri. Cu toate acestea, anumite alegeri care ne confereau un caracter distinct omenesc-iubirile de-o vară, relaţiile cu persoane pe care părinţii nu le agreează sunt mai greu de obţinut acum. În America, societatea începe să aibă un caracter deranjant, huxieyan, pe măsură ce avem relaţii, ne căsătorim şi facem copii cu persoane din aceeaşi castă, fie ea Apple, Google Blue sau Amazon.

După cum s-a anticipat, inteligenţa artificială şi automatizarea încep să elimine munca omenească mult mai multe locuri de muncă decât ne-am aşteptat. Şomajul tehnologic care se contura de multă vreme la orizont a devenit o realitate, însă nu după cum ne-am imaginat cu toţii. Eram pregătiţi ca şoferii de camion sau muncitorii să rămână fără lucru, însă previziunile s-au dovedit greşite. Am tot presupus că roboţii vor elimina slujbele slab calificate, dar se pare că construirea unor roboţi capabili de munca fizică e o sarcină mult mai grea decât ne-am imaginat, în vreme ce munca cognitivă e mult mai uşor de programat şi produs. În mod ironic, muncitorii din domeniul cunoaşterii sunt cei de care nu mai e nevoie.

Ca urmare, America şi aliaţii săi au o nevoie critică, imediată de muncitori în domeniile pe care le-am crezut pe cale de dispariţie.

Pur şi simplu nu avem suficienţi instalatori, electricieni şi tâmplari. Roboţii nu oferă aceeaşi senzaţie de omenesc de care avem nevoie, motiv pentru care maseurii, manichiuristele, cosmeticienele sau bărbierii sunt la mare căutare. împotriva automatizării se manifestă o reacţie negativă. Majoritatea oamenilor nu vor ca un robot să le pregătească băuturile sau cafeaua. Vrem să consumăm în preajma oamenilor. Faptul că am preferat disciplinele reale în favoarea celor umaniste a reprezentat o abordare întrucâtva greşită. Muncitorii sunt cei care mai au de lucru, nu informaticienii sau programatorii. Tocilarii au construit programe care i-au lăsat fără locurile de muncă.

Fără să vrea, Google, Amazon şi Apple au creat un trident în cadrul inteligenţei artificiale, puterea fiind consolidată masiv. În America şi în ţările aliate, avem la dispoziţie produse spectaculoase, dar foarte puţine alegeri. Spre exemplu, putem plăti pentru a avea ochelari Omnivision cu vedere sporită, dincolo de limitele vederii omeneşti. Doar două companii produc ochelarii: Google şi Apple. Dacă nu vă plac cum arată, sau dacă ei nu se potrivesc cu forma unică a feţei sau a urechilor, nu aveţi noroc. Amazon vinde orice vă puteţi imagina, dar produsele esenţiale zilnice reprezintă mărcile proprii ale companiei. În ţările democratice din toată lumea, există o abundenţă de mărfuri, dar varietatea şi alegerea de pe piaţă sunt strâns controlate. Deşi avem bani de cheltuială, avem o putere de cumpărare foarte redusă. Într-un mod bizar, ne aducem aminte de fosta Uniune Sovietică.

Salesforce, compania care oferă programe de relaţii cu clienţii şi cloud computing a încheiat un parteneriat cu Google, Amazon şi Apple acum mult timp pentru a construi un modul educaţional pentru DDP-uri. Acum, testarea riguroasă şi clasificarea care erau emblematice pentru

educaţia din America în anii ’80 şi 90 au devenit din nou la modă. Capacităţile noastre cognitive sunt evaluate înainte de grădiniţă, iar reuşitele şi progresele sunt monitorizate pe tot parcursul vieţii.

Metricii şi optimizarea au reprezentat dintotdeauna valorile cheie ale Salesforce, iar acum ele reprezintă valorile cheie ale educaţiei în America. Îngrijoraţi că am înlocuit înţelepciunea cu acumularea unor informaţii devenite între timp inutile, conducătorii din domeniul educaţiei au abandonat programa comună de bază în favoarea unui nou concept. Dată fiind criza forţei de muncă din America, elevii sunt împărţiţi în cadrul examenelor de admitere la grădiniţă în două categorii: filiera vocaţională sau cea de conducere. Elevii din filiera vocaţională sunt educaţi să fie agili în mai multe discipline, în timp ce elevii din filiera de conducere primesc cursuri de gândire critică şi management. Aptitudinile managementului de mijloc nu mai sunt necesare deoarece în mare parte activitatea managerilor de mijloc şi a lucrătorilor din domeniul cunoaşterii este îndeplinită acum de inteligenţa artificială.

Dat fiind şomajul apărut în domenii unde nu ne aşteptam, şi infracţionalitatea a crescut-însa nu din motivele la care v-aţi fi aşteptat. Programele de ordine publică bazate pe inteligenţă artificială nu au funcţionat cum trebuie, aşa că statisticile privind criminalitatea nu reflectă în mod precis lumea reală. Algoritmii construiţi de triburile inteligenţei artificiale şi instruiţi pe baza unui set limitat de date nu au învăţat niciodată cum să identifice şi să clasifice o persoană cu tulburări de geno persoană care nu se identifică nici ca bărbat, nici ca femeie şi care poate avea un aspect androgin sau o barbă dar şi gene false. Ca urmare, sute de oameni care nu îndeplinesc criteriile unui anume gen sunt acuzaţi în mod eronat în fiecare zi că fură identitatea altora: când încearcă să plătească prin recunoaştere facială, când se plimbă prin birou sau când încearcă să efectueze un apel video. Pentru moment,

singura soluţie este să încerce să se conformeze unui anumit gen în timpul anumitor tranzacţii. Ei/ele trebuie să-şi pună o anumită perucă specifică genului, sau sa de demachieze pentru a deveni temporar un el/ea în ochii inteligenţei artificiale care procesează datele vizuale. E un memento umilitor şi public al faptului că diversitatea nu a contat niciodată suficient de mult încât să reparăm un sistem defect.

Inteligenţa artificială conferă Google, Apple şi Amazon o putere economică uriaşăiar Chinei o putere militară şi geopolitică de neimaginat. La finalul anilor ’2030, ne-am dat seama că inteligenţa artificială s-a dezvoltat de-a lungul unor traiectorii paralele, sprijinind capitalismul în Occident, şi varianta chineză a comunismului în Asia, Africa şi America latină. America şi aliaţii săi, care sărbătoreau odinioară reuşitele Mafiei-G, trăiesc acum într-un sistem totalitar bazat pe inteligenţă artificială. Cetăţenii chinezi şi ai ţărilor care beneficiază de investiţiile directe şi infrastructura chineză constată că şi ei trăiesc sub un aparat omniprezent de pedepse şi recompense bazat pe inteligenţa artificială.

2049: Graniţe biometrice şi avorturi cu nano-boţi

Mafia-G a devenit grupul GAA: Google, Apple şi Amazon. Facebook a fost prima dintre companii care şi-a declarat falimentul. Rămăşiţele Microsoft şi IBM au fost achiziţionate de Google.

Sărbătorim 100 de ani de la Revoluţia Comunistă din China şi proclamarea Republicii Populare Chineze (RPC) de către Mao Zedung. Ample festivităţi sunt anunţate în toate ţările satelit ale Chinei pentru onora memoria defunctului Xi Jinping şi a sărbători ascensiunea dinastiei Rengong Zhineng (Inteligenţa Artificială).

Toată omenirea este acum înconjurată de sisteme de inteligenţă artificială, care ar fi trebuit să ne ajute să ducem vieţi mai libere şi mai fericite. Încă de la început, triburile inteligenţei artificiale din Statele Unite au spus că vor să trăim cât mai bine, să avem activităţi creative şi să

colaborăm în privinţa celor mai mari provocări ale omenirii. A fost un ideal utopic născut în mediul izolat din Silicon Valley, iar creatorii acestui ideal pierduseră complet orice legătură cu lumea din afară.

Toate aceste sisteme au fost create să ne facă viaţa mai uşoară dar au reuşit în schimb să ne facă mai leneşi. Ele ne-au subminat senzaţia de productivitate, de a avea un scop. Ne bazăm pe aceste sisteme pentru a decide în locul nostru. Ne resemnăm la nişte alegeri limitate. Trecem prin etapele preprogramate ale traiului cotidian, optimizate de inteligenţa artificială generală pentru toţi locuitorii planetei.

Multe sisteme LAG au evoluat pentru a concura, nu pentru a colabora. Atacurile Chinei asupra interfeţelor aparatelor de acum două decenii ne par atât de simpliste, de benigne. Trăiţi într-o închisoare bazată pe inteligenţa artificială pe care v-aţi construit-o singuri. Pierdeţi controlul constant asupra cuptorului, şifonierului sau băii, şi nici nu vă mai deranjaţi să vă mai luptaţi. Nu mai are sens. Reacţia normală, vi s-a spus, este să aşteptaţi ca problema să se rezolve. Casele Google Green şi Apple pot achiziţiona un abonament suplimentar de top, care permite teoretic accesul la o IAG de reparaţie care să suprascrie codul greşit-însa IAG-urile sunt blocate într-o buclă a auto-perfecţionării. Toţi banii din lume nu pot repara o casă afectată de o defecţiune sistemică.

Concentrarea averilor a permis GAA să obţină rezultate uimitoare în domeniul sănătăţii. Google a fost prima companie care a lansat un proiect pilot cu roboţi microscopici, injectabili, capabili să livreze medicamentele doar unei anume părţi a corpului sau să asiste în micro-chirurgie. Nanoboţii sunt fabricaţi acum sub multe forme diferite. Spre exemplu, există un robot autonom, molecular, construit dintr-o singură fâşie de ADN care tratează interiorul corpului omenesc ca pe un depozit de distribuţie. Acest nanobot se poate plimba, luând molecule dintr-un loc şi depunându-le în alte locuri stabilite. Un alt

tip de nanobot, care se deplasează pe baza unor bule de aer, poate să administreze cantităţi microscopice de medicamente fără să provoace rău. Apariţia acestor nanoboţi pe piaţă, care partajează informaţii cu DDP-urile noastre, a înlocuit terapiile şi medicaţiile generalizate, tratând suferinţele noastre distincte fără să provoace efecte secundare.

Acum că atât Amazon cât şi Apple oferă servicii medicale personalizate, majoritatea oamenilor s-au injectat de bunăvoie cu nano-boţi organici. Chiar şi familiile Amazon au acces printr-un program subvenţionat aprobat de guvernul Statelor Unite. Nanoboţii continuă să ne monitorizeze şi să ne trateze, astfel încât speranţa de viaţă a unui american de rând a crescut de la 76, 1 ani în 2019 la 99, 7 ani în zilele noastre.184

Nu a durat prea mult până când am descoperit şi potenţialele efecte negative ale IAG injectabile. Nanoboţii au făcut exact ceea ce au dorit creatorii lor. Ei s-au comportat imprevizibil şi au început să înveţe. Dacă ne gândim în urmă, construirea şi instruirea sistemelor de inteligenţă artificială pentru a face lucruri la care nu ne-am fi gândit niciodată a fost un obiectiv principal al triburilor inteligenţei artificiale. Aceasta a fost cheia rezolvării problemelor spinoase pe care omenirea nu le-ar fi putut confrunta de una singură. Când Alpha Go Zero a luat decizii strategice autonome în urmă cu decenii, am considerat că inteligenţa artificială atingea noi culmi. În interiorul corpurilor noastre, însă, nanoboţii şi LAG-urile care le controlează se îmbunătăţesc singure şi au mai multă putere de decizie decât ne-am fi dorit.

Avem acum o nouă *himeră economică* a oamenilor. Casele Apple şi Google Green pot accesa super-puteri, obţinând o cogniţie superioară, un miros extrasenzorial şi o atingere îmbunătăţită.

Casele Google Blue, Yellow şi Amazon doar că nu au acces la aceste îmbunătăţiri ci au şi anumite restricţii biologice. Când o persoană rămâne însărcinată, LAG-urile

rulează constant modele predictive pentru a stabili sănătatea şi viabilitatea fetusului. Ceea ce nimeni nu anticipase era ca LAG-urile să-şi ducă obiectivele la extrem. Din cauză că obiectivul programat era de a sprijini oamenii să aducă pe lume fetuşi sănătoşi, IAG-urile au început să caute anomalii aleţ esutului la fat. Dacă se detecta o astfel de anomalie, IAG avorta automat fătul, fără să ofere părintelui opţiunea de a alege dacă păstrează copilul.

În mod similar, nanoboţii monitorizează procesul de îmbătrânire, calculând când devine continuarea vieţii mai dureroasă decât moartea. Odată ce începeţi să aveţi nevoie de îngrijire la domiciliu şi deveniţi o povară asupra resurselor statului, IAG-urile intervin, inducând o moarte confortabilă, pentru ca nici voi, şi nici rudele voastră, să nu trebuiască să decidă când e vremea să muriţi.

Legile ţărilor GAA au fost depăşite o dată ce LAG-urile s-au îmbunătăţit şi au putut crea mecanismele care stabilesc cine trăieşte şi cine moare. Guvernele s-au grăbit să creeze un cadru juridic şi reglementări care nu ajută însă la nimic. Interzicerea nanoboţilor ar presupune revenirea la medicina tradiţională, iar marile companii farmaceutice care fabricau medicamentele de care am fi avut nevoie au dispărut. Chiar şi cele mai optimiste proiecţii ne arată că repornirea fostelor sisteme sanitare ar presupune un efort de un deceniu sau chiar mai bineiar, între timp, milioane de oameni ar suferi de multiple boli.

Cercetătorii au dezvoltat în schimb un nou tip de nanobot bazat pe LAG care poate controla alţi nanoboţi din interiorul corpului-imitând felul în care leucocitele luptă cu un virus. La fel ca inteligenţa artificială, ideea a fost inspirată de biologia omului. Pe măsură ce corpurile noastre se luptă cu nanoboţii LAG indezirabili, simptomele sunt mult mai neplăcute decât cele ale unei gripe, şi mult mai periculoase.

Marile companii sunt conduse acum de directori generali reprezentaţi de inteligenţa artificială sunt cei care

calculează riscurile şi oportunităţile strategice. Directorii umani lucrează alături de aceste entităţi, reprezentând „faţa” companiei. Companiile de dimensiuni mici şi mediirestaurantele, service-urile auto, saloanele cosmetice sunt toate partenere ale uneia din grupul GAA. Pe lângă DDP-urile personale şi ale caselor, fiecare firmă şi ONG este acum înregistrată într-un Dosar de Date ale Organizaţiei.

Cu toate acestea, foarte mulţi oameni din America şi din ţările aliate nu au un loc de muncă. În absenţa unei plase sociale de siguranţă, economiile occidentale se află într-un declin prelungit, deoarece nu s-au putut redresa după valurile neanticipate de şomaj tehnologic. Astfel s-au creat vulnerabilităţişi o fereastră pentru investiţiile chineze. În curând, politicienii sunt forţaţi să aleagă între idealurile democratice şi viabilitatea economicão decizie deosebit de grea pentru persoane care se confruntă cu alegerile şi presiunea de a rezolva problemele domestice urgente.

Ca răspuns, Statele Unite încearcă să limiteze expansiunea Chinei prin blocade comerciale, sancţiuni secundare şi alte tactici diplomatice. America constată însă că nu mai are influenţa geopolitică de odinioară. Liderii americani au petrecut prea mult timp deliberând în loc să acţioneze în privinţa Chinei. Ei au călătorit prea puţin în America Latină, Africa şi Asia de Sud-Est. Ei nu şi-au câştigat niciodată încrederea, favoarea şi prietenia omologilor lor din străinătate.

Demersurile chineze în domeniul inteligenţei artificiale devin din ce în ce mai importante. Scorul de armonie socială este folosit în peste 100 de ţări, înlocuind documentele tradiţionale de călătorie. China excelează în construcţia zidurilor, iar Marele Zid al Inteligenţei Artificiale nu face excepţie de la această regulă. Zidul oferă atât o barieră de protecţie împotriva străinilor, dar şi o metodă prin care datele tuturor pot fi extrase şi analizate. Cei care au un scor al armoniei sociale suficient

de ridicat primesc acces nelimitat (dar monitorizat, desigur) prin Marele Zid al Inteligenţei Artificiale către oricare dintre ţările conectate la reţeaua Chinei. China a înfiinţat graniţe biometrice cu recunoaştere facială pentru a putea stabili cine poate intra sau ieşi. Nu mai există poliţie de frontieră, şi nici paşapoarte care să fie ştampilate.

La graniţa de sud a Statelor Unite există acum un zid. El este un zid de senzori şi a fost construit pe pământ mexican de către chinezi, pentru a ne izola. Deoarece americanii nu au acces la sistemul de credite sociale, vi se refuză intrarea în fostele destinaţii preferate de vacanţă: Bahamas, Jamaica, Cancún, Playa del Carmen, Cozumel, Costa Rica şi Aruba. Dacă încercaţi să treceţi ilegal o frontieră biometrică, o IAG emite un atac sonic care provoacă greaţă, contuzii, sângerări auriculare şi stres psihologic.

Americanii şi aliaţii noştri sunt izolaţi de prietenii şi rudele dinţările satelit ale Chinei, deoarece PCC controlează toată infrastructura reţelelor din aceste ţări. Dacă aveţi nevoie să căutaţi pe cineva dintr-o ţară aflată sub controlul PCC, trebuie să treceţi prin China ca intermediar, ştiind că toate cuvintele sunt ascultate.

Grupul GAA a format în cele din urmă o coaliţie cu guvernul Statelor Unite şi aliaţii rămaşi. Date fiind restricţiile economice şi de călătorie impuse de China, există puţine resurse financiare prin care să poată fi găsită o soluţie acceptabilă. Se ia decizia de a dezvolta o IAG care să rezolve problema Chinei pentru noi. Sistemul vede însă doar două soluţii: să cedăm Chinei, sau să reducem gradual populaţia globului.

2069: Anihilarea digitală în timp ce China se preocupa de planificarea pe termen lung şi o strategie naţională în domeniul inteligenţei artificiale, Statele Unite erau preocupate de aparate şi dolari.

China nu mai are acum nevoie de Statele Unite ca partener comercial şi nu are nevoie de drepturile noastre

de autor. China a construit o reţea de peste 150 de ţări care operează pe baza principiilor directoare ale politicii globale chineze. La schimb pentru obedienţa lor, aceste ţări primesc acces la reţea, capacitatea de a face comerţ şi un sistem financiar stabil sprijinit de Beijing. Cetăţenii lor pot călători liber în spaţiul chinez, cu condiţia să aibă un scor social suficient de ridicat.

Libertatea de mişcare – un drept pe care americanii considerau că li se cuvinenu ne-a lipsit niciodată mai mult. Motivul e că America trece, la fel ca multe alte ţări, printr-o criză a populaţiei. Pe planeta noastră trăiesc peste 10 miliarde de oameni. Se nasc prea mulţi oameni, şi prea repede, şi cu toate acestea insistăm să ne prelungim viaţa până dincolo de 120 de ani.

Populaţia planetei constituie o problemă pentru că nu am combătut suficient de repede încălzirea globală, nici măcar după ce China a preluat răspunderea sustenabilităţii şi a protecţiei mediului. Două treimi din terenurile arabile au fost pierdute. Deşi facem eforturi ca să construim ferme subterane în America, nu putem creşte mâncarea suficient de repede pentru a hrăni populaţia. Sancţiunile globale au blocat rutele comerciale şi ne-au izolat de naţiunile care exportă mâncare, dar şi China şi sateliţii săi au dificultăţi.

Într-o bună zi, familiile Apple suferă de o boală misterioasă. DDP-urile lor prezintă o anomalie, fără să ofere amănunte sau detalii concrete. La început, credem că această ultima versiune de nanoboţi e defectă, aşa că managerii de produs se grăbesc să ofere variante de IAG care să repare deficienţele. Boala loveşte apoi casele Googlenu doar în America, ci şi în toate locuinţele din afara spaţiului chinez. Boala misterioasă se agravează rapid.

China a construit o SIA care are un singur scop: exterminarea populaţiilor Americii şi a aliaţilor acesteia. Ţările din arealul chinezesc au nevoie de resursele rămase ale planete, iar Beijingul a calculat că unica metodă de a

supravieţui este de a ne lua aceste resurse.

Ceea ce urmează este mai grav decât orice bombă inventată vreodată. Bombele au efecte imediat şi distrugător. Anihilarea prin inteligenţa artificială este înceată şi ireversibilă. Rămâneţi neajutoraţi în timp ce corpurile copiilor vă mor în braţe. Vedeţi cum colegii se prăbuşesc la birou. Simţiţi o durere ascuţită. Vă ia ameţeala. Respiraţi o ultimă respiraţie, repede, superficial.

Iată sfârşitul Americii.

Iată sfârşitul aliaţilor Americii.

Iată sfârşitul democraţiei.

Dinastia Rengong Zhineng preia puterea, brutal, ireversibil şi absolut.

Există în prezent semnale care validează fiecare dintre aceste trei scenarii. Trebuie să alegem o variantă acum. Voi trebuie să alegeţi. Vă cer să alegeţi scenariul optimist, şi să construiţi un viitor mai bun pentru inteligenţa artificială, şi pentru omenire.

PARTEA A III-A:

Rezolvarea problemelor

CAPITOLUL OPT

PIETRICELE ŞI BOLOVANI: CUM SĂ REPARĂM VIITORUL INTELIGENŢEI ARTIFICIALE

Concluziile ultimului capitol ar putea părea extreme şi improbabile. Există însă deja semnale care ne spun că dacă nu alegem un viitor în care cei Nouă Titani să fie stimulaţi să colaboreze în interesul omenirii, este foarte probabil să ajungem să trăim într-o lume care să semene cu dinastia Rengong Zhineng.

Cred că scenariul optimistsau o variantă apropiată de acesta e la îndemână. Este posibil ca inteligenţa artificială să-şi îndeplinească scopul superior şi potenţialul, spre beneficiul triburilor inteligenţei artificiale dar şi al tuturor oamenilor. Pe măsură ce evoluează, inteligenţa artificială poate deservi atât oamenii din China şi din Statele Unite, cât şi toţi aliaţii noştri. Ea ne poate ajuta să ducem o viaţă mai sănătoasă, să reducem ecarturile economice şi să ne

sporească siguranţa, pe stradă sau în casă. Inteligenţa artificială ne poate oferi puterea de a descoperi tainele întrebărilor celor mai profunde ale omenirii, cum ar fi originile şi modul în care a apărut viaţa. În cadrul acestui proces, inteligenţa artificială ne va putea ului şi distra, creând lumi virtuale pe care nu ni le-am fi putut niciodată imagina, scriind cântece care să ne inspire şi creând noi experienţe distractive şi care să ne împlinească. Nu vom putea avea însă niciunul dintre aceste beneficii fără planificare, fără un angajament de a pune osul la muncă şi fără un exemplu curajos al tuturor celor implicaţi în grupurile care dezvoltă inteligenţa artificială.

Tehnologiile sigure şi benefice nu sunt rezultatul speranţelor şi al hazardului. Ele sunt produse ca urmare a contribuţiilor unor lideri curajoşi, şi a unor colaborări dedicate, de lungă durată. Cei Nouă Titani se află sub o presiune imensă – din partea sistemului financiar în Statele Unite, respectiv a Beijingului în Chinade a îndeplini aşteptări pe termen scurt, chiar dacă periclitează astfel viitorul nostru. Trebuie să emancipăm cei Nouă Titani şi să le insuflăm curajul de a schimba traiectoria inteligenţei artificiale, deoarece în absenţa unui sprijin deplin din partea noastră, aceste companii nu pot şi nu vor dori să reuşească pe cont propriu.

Vânt Cerf, unul dintre cei care a elaborat primele protocoale şi arhitectura internetului modern, foloseşte o parabolă pentru a explica de ce conducerea curajoasă este esenţială în epoca tehnologiilor emergente, precum inteligenţa artificială.185 Imaginaţi-vă că trăiţi într-o comunitate minusculă într-o vale înconjurată de munţi. În vârful unui munte îndepărtat se află un bolovan uriaş. A stat acolo de mult timp şi nu s-a mişcat niciodată, aşa că din punctul de vedere al comunităţii voastre, e doar parte din peisaj, într-o bună zi, observaţi că bolovanul uriaş pare instabilcă se pregăteşte să se rostogolească la vale, cu o viteză şi o putere din ce în ce mai mare, şi că va distruge comunitatea voastră şi pe toţi membrii ei. De fapt, vă daţi

seama că poate toată viaţa aţi ignorat mişcarea lui. Acel gigantic bolovan s-a mişcat dintotdeauna, puţin câte puţin, însă nu aţi acordat atenţia voastră deplină schimbărilor subtile, minuscule, de fiecare zi: o mică schimbare a umbrei pe care o aruncă, distanţa dintre el şi următorul munte, sunetul aproape imperceptibil pe care-l produce în timp ce striveşte pământul de sub el. Vă daţi seama că o singură persoană nu poate urca pe munte şi să oprească bolovanul gigantic de una singură. Sunteţi prea mic, şi bolovanul e prea mare.

Vă daţi apoi seama că dacă puteţi găsi o pietricică pe care să o plasaţi într-un loc bun, ea va încetini mişcarea bolovanului şi îi va devia un pic traiectoria. O singură pietricică nu va împiedica bolovanul să distrugă satul, aşa că cereţi întregii comunităţi să vi se alăture. Cu pietricelele în mână, fiecare persoană urcă muntele şi se pregăteşte pentru venirea bolovanului există colaborare, şi comunicare, şi un plan de a confrunta bolovanul pe măsură ce coboară. Oamenii şi pietricelele lornu un bolovan mai marereuşesc să facă diferenţa.

Ceea ce urmează este o serie de pietricele. Voi începe foarte larg, subliniind motivele pentru constituirea unei comisii la nivel global care să supervizeze parcursul inteligenţei artificiale şi nevoia noastră imediată de a avea norme şi standarde. Voi explica apoi ce schimbări vor trebui luate de către guvernele Chinei şi Statelor Unite. După aceea, voi arunca o privire mai în amănunt asupra nevoii celor Nouă Titani de a-şi reforma practicile. Mă voi concentra apoi doar asupra triburilor inteligenţei artificiale şi a universităţilor în care acestea se formează şi voi explica în amănunt ce schimbări trebuie făcute chiar acum. În cele din urmă, voi explica rolul pe care şi voi îl puteţi juca în modelarea viitorului inteligenţei artificiale.

Viitorul în care cu toţii ne vom *dori* să trăim nu se va împlini de la sine şi complet. Trebuie să fim curajoşi. Trebuie să ne asumăm răspunderea pentru acţiunile noastre.

Schimbarea sistemică globală: argumentul pentru înfiinţarea GAIA

În scenariul optimist, un amestec divers de lideri din economiile cele mai avansate îşi unesc eforturile cu Mafia G pentru a înfiinţa Alianţa Globală de Augmentare a Inteligenţei, sau GAIA. Acest corp internaţional include cercetători din domeniul inteligenţei artificiale, sociologi, economişti, teoreticieni ai jocurilor, futurologi, politicieni din toate statele membre. Membrii GAIA reflectă o diversitate socioeconomică, de gen, rasă, religie, politică şi sexuală. Ei sunt de acord să faciliteze şi să coopereze în iniţiative comune şi politici legate de inteligenţa artificială, reuşind cu timpul să exercite suficientă influenţă şi control încât să împiedice o apocalipsă-fie ea datorată LAG, SLA sau folosirii inteligenţei artificiale în China pentru a oprima cetăţenii.

Cea mai bună modalitate de a îndeplini schimbarea sistemică este de a vedea înfiinţarea GAIA cât de curând posibil, iar aceasta ar trebui să aibă sediul pe un teren neutru, în apropierea unui centru de inteligenţă artificială. Cel mai bun loc în care GAIA ar trebui să-şi aibă sediul este în Montreal, Canada. Mai întâi, Montreal găzduieşte multe laboratoare de cercetare şi cercetători în domeniul învăţării profunde. Dacă presupunem că tranziţia de la LAÎ la IAG va include învăţarea profundă şi reţele neurale profunde, rezultă că GAIA ar trebui să îşi aibă sediul în apropiere de locul unde se desfăşoară această activitate. În al doilea rând, în regimul prim-ministrului Justin Trudeau, guvernul canadian a alocat deja personalul şi resursele financiare pentru a explora viitorul inteligenţei artificiale. În anii ’2017 şi 2018, Trudeau nu s-a rezumat să vorbească despre inteligenţa artificială, ci a poziţionat Canada pentru a putea contribui la modelarea reglementărilor şi principiilor care stabilesc evoluţia inteligenţei artificiale. În al treilea rând, Canada reprezintă un teritoriu geopolitic neutru pentru inteligenţa artificială este departe atât de Silicon Valley cât şi de Beijing.

Poate părea imposibil ca guvernele lumii să se unească în jurul unei cauze, date fiind tensiunile politice şi instabilitatea geopolitică din ultimii ani. Există însă şi un precedent. La finalul celui de-al doilea război mondial, când tensiunile erau încă ridicate, sute de delegaţi din toate ţările aliate s-au adunat la Bretton Woods, în statul New Hampshire, pentru a construi structurile financiare care au permis economiei globale să meargă mai departe. Această colaborare a fost centrată pe oameni-ea a rezultat într-un viitor în care oamenii şi ţările au putut reconstrui şi să caute prosperitatea. Ţările GAIA ar trebui să colaboreze în domeniul cadrelor de referinţă, standardelor şi bunelor practici în domeniul inteligenţei artificiale. Deşi este improbabil ca China să se alăture, liderii PCC şi ai grupului BAT ar trebui să fie invitaţi.

Înainte de orice, GAIA va trebui să stabilească o modalitate prin care să garanteze drepturile elementare ale omului în epoca inteligenţei artificiale. Când vorbim despre inteligenţă artificială şi etică, avem tendinţa să ne gândim la cele trei legi ale roboticii ale lui Isaac Asimov, pe care le-a publicat într-o nuvelă din 1942 intitulată „Runaround”. 186 Era o poveste despre un computer umanoid, nu despre inteligenţa artificială, însă cu toate acestea aceste legi ne-au inspirat gândirea etică atât de mulţi ani mai târziu. După cum am amintit în Capitolul 1, regulile lui Asimov sunt sunt: (1) Un robot nu poate răni un om, sau prin inacţiune să permită ca un om să sufere. (2) Un robot trebuie să asculte ordinele pe care i le dau oamenii, mai puţin în cazurile în care aceste ordine s-ar afla în conflict cu Prima Lege. (3) Un robot trebuie să-şi protejeze existenţa atât timp cât această protecţie nu intră în conflict cu Prima sau a Doua Lege.

Când Asimov a publicat ulterior o colecţie de nuvele într-o carte intitulată I, *Robot*, el a adăugat o lege „Zero”, care să le supună pe toate celelalte: (O) roboţii nu pot face rău oamenilor. Asimov a fost un scriitor talentat, clarvăzător-dar legile sale ale roboticii sunt prea generale

pentru a servi drept principii directoare pentru viitorul inteligenţei artificiale.

GAIA ar trebui în schimb să creeze un nou contract social între cetăţeni şi cei Nouă Titani (definiţi în mod extins ca fiind Mafia G şi grupul BAT precum şi partenerii, investitorii şi filialele acestora). Contractul trebuie să se bazeze pe încredere şi colaborare. Membrii GAIA ar trebui să fie de acord că inteligenţa artificială trebuie să emancipeze cât mai mulţi oameni din jurul lumii. Cei Nouă Titani ar trebui să acorde prioritate drepturilor omului şi nu ar trebui să ne considere nişte resurse care să fie exploatate pentru profit sau scopuri politice. Prosperitatea economică promisă de inteligenţa artificială şi de cei Nouă Titani ar trebui să sporească considerabil bunăstarea tuturor.

Rezultă aşadar că Dosarele noastre de Date Personale ar trebui să poată fi portate între operatori şi ar trebui să fie proprietatea noastră, nu a unor companii, sau consorţii, sau naţiuni. GAIA ar putea începe încă de azi să studieze cum se poate implementa această măsură, pentru că DDP-urile despre care este vorba în scenarii există deja în zilele noastre, în formă incipientă. Ele se numesc „Informaţii Personale Identificabile” sau IPI. Aceste IPI individuale sunt cele pe baza cărora funcţionează aplicaţiile din telefoanele noastre inteligente, reţelele de publicitate din site-uri, precum şi recomandările care ne sugestionează pe ecrane. IPI sunt transmise unor sisteme folosite pentru a ne identifica şi localiza. Felul în care sunt folosite depinde în întregime de capriciile companiilor şi agenţiilor guvernamentale care le accesează.

Înainte de a genera un nou contract social, GAIA trebuie să decidă cum pot fi folosite DDP-urile pentru a ajuta algoritmii de învăţare ai maşinilor, şi trebuie să definească ce anume constituie valorile de bază în epoca automatizării. Definirea clară a valorilor are o importanţă esenţială pentru că aceste valori sunt cele care vor fi în cele din urmă programate în datele de instruire, datele din

lumea reală, sistemele de învăţare şi aplicaţiile care alcătuiesc ecosistemul inteligenţei artificiale.

Pentru a putea cataloga valorile noastre de bază, GAIA ar trebui să creeze un Atlas al Valorilor Omenirii, care să definească valorile noastre unice în funcţie de diferite ţări şi culturi. Acest atlas nu ar trebui să fie unul static. Deoarece valorile se schimbă cu timpul, atlasul ar trebui să fie actualizat de ţările membre. Putem căuta precedentul în biologie: Atlasul Celulelor Umane reprezintă o colaborare globală a comunităţii ştiinţifice, care include mii de experţi din diferite domenii (genetică, inteligenţă artificială, programare, vizualizarea datelor, medicină, chimie şi biologie). 187 Proiectul cataloghează fiecare tip de celulă din corpul uman, cartografiind tipologiile celulare cu locul unde se găsesc, urmărind istoria celulelor pe măsură ce evoluează şi surprinzând caracteristicile celulelor pe durata vieţii. Acest demers-costisitor, complicat, cronofag şi continuu-va permite cercetătorilor să facă progrese semnificative, iar el este posibil doar graţie unei colaborări substanţiale, la nivel mondial. Ar trebui să avem un atlas asemănător al valorilor umane, în care să fie implicaţi profesori, antropologi, sociologi, psihologi dar şi oameni obişnuiţi. Crearea Atlasului Valorilor Umane ar fi un demers scump, problematic şi greoişi ne-am lovi de multe contradicţii, deoarece valorile unor culturi ar fi contrare cu ale altora. Cu toate acestea, în absenţa unui cadru şi a unui set de standarde elementare, le cerem celor Nouă Titani şi triburilor inteligenţei artificiale să facă ceva imposibil să ia în considerare toate perspectivele noastre şi toate rezultatele posibile în ce priveşte grupuri distincte din cadrul societăţii, din fiecare ţară din lume.

GAIA ar trebui să ia în considerare o structură de drepturi juridice care să pună în balanţă libertăţile individuale şi binele comun. Ar fi de preferat ca structura dezvoltată să reprezinte bine idealurile dar să ofere o interpretare mai flexibilă, pe măsură ce inteligenţa

artificială evoluează. Organizaţiile membre ar trebui să poată demonstra că respectă această structură sau să fie ameninţate cu excluderea din GAIA. Orice structură ar trebui să includă următoarele principii:

1. Omenirea ar trebui să se afle întotdeauna în centrul evoluţiei inteligenţei artificiale.

2. Sistemele de inteligenţă artificială ar trebui să fie sigure. Ar trebui să putem verifica singuri cât de sigure sunt.

3. Cei Nouă Titani – inclusiv investitorii, angajaţii şi guvernele pentru care lucrează trebuie să optimizeze siguranţa în favoarea vitezei. Orice echipă care lucrează la un sistem de inteligenţă artificială – chiar şi cele din afara celor Nouă Titanitrebuie să acorde prioritate siguranţei, şi nu vitezei. Siguranţa trebuie să poată fi demonstrată şi confirmată de observatori independenţi.

4. Dacă un sistem de inteligenţă artificială face rău, el trebuie să poată da socoteală privind erorile fiind necesar un proces decizional prin care pagubele să poată fi analizate şi combătute.

5. Inteligenţa artificială trebuie să fie explicabilă. Sistemele ar trebui să deţină un soi de etichetă cu informaţii nutritive, prin care să explice ce fel de date de instruire au fost folosite, ce procese de învăţare, datele din lumea reală folosite în aplicaţii şi rezultatele anticipate. În ce priveşte sistemele sensibile sau protejate de drepturi de autor, transparenţa unei inteligenţe artificiale ar trebui să poată fi evaluată şi verificată de către terţi.

6. Toţi membrii ecosistemului inteligenţei artificiale-persoanele care angajează, conduc şi supraveghează cei Nouă Titani, firmele noi (antreprenori, acceleratoare de afaceri); investitori (firme private de investiţii, fonduri mutuale şi acţionari individuali); profesori şi studenţi post-universitari şi orice alte persoane care lucrează în domeniul inteligenţei artificiale-trebuie să recunoască faptul că iau tot timpul decizii cu caracter etic. Ei trebuie să fie gata să explice toate deciziile pe care le-au luat în

cadrul procesului de dezvoltare, testare şi implementare.

7. Atlasul Valorilor Umane trebuie respectat în cadrul tuturor proiectelor de inteligenţă artificială. Chiar şi aplicaţiile înguste ale inteligenţei artificiale trebuie să demonstreze faptul că respectă Atlasul.

8. Trebuie să existe un cod de conduită publicat şi uşor de găsit care să guverneze persoanele care lucrează în inteligenţa artificială în etapele de proiectare, dezvoltare şi implementare. Codul de conduită trebuie să se aplice şi investitorilor.

9. Toţi oamenii ar trebui să aibă dreptul să pună întrebări sistemelor de inteligenţă artificială. Care este adevăratul scop al unei inteligenţe artificiale, ce date foloseşte, cum ajunge la concluzii şi cine poate vedea rezultateleaccesul la aceste informaţii ar trebui să fie complet transparent şi într-un format standardizat.

10. Termenii de operare ai unei aplicaţii care include inteligenţa artificială-sau ai oricărui serviciu care foloseşte inteligenţa artificială – ar trebui să fie formulaţi într-un limbaj care este suficient de simplu încât un elev de clasa a treia să-l poată înţelege. Aceşti termeni ar trebui traduşi în toate limbile încă de la lansarea aplicaţiei.

11. DDP-urile ar trebui să fie opţionale şi dezvoltate folosind un format standardizat, ar trebui să fie interoperabile iar oamenii ar trebui să deţină în totalitate dreptul de proprietate şi de acces. Dacă aceste DDP-uri ar deveni transmisibile, oamenii ar trebui să poată alege ce drepturi de acces vor să cedeze şi cum sunt folosite datele lor.

12. DDP-urile ar trebui descentralizate pe cât posibil, asigurând că nicio parte nu deţine controlul complet asupra lor. Grupul de lucru care creează formatul DDP-urilor ar trebui că în ciudă şi experţi juridici şi din afara sistemului juridic: hackeri buni, conducători ai societăţii civile, agenţi guvernamentali, depozitari independenţi de date, eticieni şi alţi profesionişti din afara celor Nouă Titani.

13. Pe cât posibil, DDP-urile ar trebui protejate pentru a nu spori puterea regimurilor totalitare.

14. Trebuie să existe un sistem prin care să se poată răspunde în faţa publicului şi o metodă simplă prin care oamenii să primească răspunsuri la întrebările legate de datele proprii şi felul în care ele sunt extrase, rafinate şi folosite în sistemele de inteligenţă artificială.

15. Toate datele ar trebui să fie tratate în mod egal şi echitabil, indiferent de naţionalitate, rasă, religie, identitate sexuală, gen, afinităţi politice sau alte convingeri proprii.

Membrii GAIA ar trebui să fie gata să accepte inspecţii prin sondaj ale altor membri sau ale unei agenţii din cadrul GAIA pentru a asigura că acest cadru este respectat în totalitate. Toate amănuntele-cum ar fi cele legate de un sistem de audit şi felul în care acesta funcţionează în lumea realăvor trebui constant revizuite şi îmbunătăţite, pentru a putea ţine pasul cu evoluţiile din domeniul inteligenţei artificiale. Acest proces ar încetini cu siguranţă anumite progrese, şi totul ar fi intenţionat.

Organizaţiile şi ţările membre ar trebui să colaboreze şi să împartă descoperirile, printre care se numără şi vulnerabilităţile sau riscurile de securitate. Astfel, membrii GAIA ar putea avea un atu asupra agenţilor malefici care ar putea încerca să dezvolte capacităţi nocive ale inteligenţei artificiale, cum ar fi sisteme autonome de hacking. Deşi pare improbabil ca cei Nouă Titani să fie dornici să partajeze secretele comerciale, avem şi aici un precedent: Organizaţia Mondială a Sănătăţii coordonează reacţiile la nivel global în momentele de criză, în timp ce un grup denumit Centrul pentru Securitate Cibernetică Avansată mobilizează forţele de ordine, cercetătorii universitari şi departamentele guvernamentale în timpul unor atacuri cibernetice. Această coordonare ar permite membrilor GAIA să dezvolte o serie de inteligenţe artificiale santinelă, care ar identifica mai întâi dacă un sistem de inteligenţă artificială se comportă aşa cum a fost

prevăzut-nu doar codul, ci şi felul în care foloseşte datele noastre şi cum interacţionează cu aparatele în care este instalat. Inteligenţele artificiale de tip santinelă ar dovedi în mod formal că sistemele de inteligenţă artificială îşi fac treaba aşa cum ar trebuie, iar pe măsură ce ecosistemul inteligenţei artificiale se maturizează către IAG, orice schimbă efectuate autonom care ar putea modifica obiectivele existente ale unui sistem ar fi raportate înainte ca autoîmbunătăţirea să poată fi efectuată. Spre exemplu, o inteligenţă artificială santinelă-un sistem creat pentru a monitoriza şi raporta activitatea altor inteligenţe artificialear putea revizui rezultatele într-o reţea antagonică generativă, amintită în capitolele anterioare, şi să se asigure că sistemul funcţionează aşa cum a fost proiectat. Odată ce trecem de la L-AI la LAG, sistemele de tip santinelă ar continua să raporteze şi să verificedar nu ar fi programate să acţioneze autonom.

Odată ce ne apropiem de IAG, cei Nouă Titani şi toţi membrii ecosistemului inteligenţei artificiale ar trebui să fie de acord să restricţioneze inteligenţa artificială într-un mediu de test şi să simuleze riscurile înainte de a o implementa în lumea reală. Sugestia mea este radical diferită de varianta actuală a testării produselor, care caută în principal să stabilească dacă un sistem operează aşa cum a fost proiectat. Deoarece nu putem cunoaşte toate modurile posibile în care o tehnologie ar putea evolua sau să fie refolosită în lumea reală înainte de a o lansa, trebuie să rulăm simulări tehnice şi să catalogăm riscurile pentru a vedea consecinţele asupra libertăţilor economice, geopolitice şi personale. Inteligenţa artificială ar trebui să fie limitată până când vom putea şti că beneficiile sunt mai mari decât posibilele consecinţe negative, sau până când vom găsi o metodă de a contracara riscurile. Aceasta înseamnă să le permitem celor Nouă Titani să-şi ducă la capăt cercetarea fără ameninţarea constantă a unor raportări financiare şi prezentări de produse.

Schimbări la nivel guvernamental: argumente pentru o reorientare a Statelor Unite şi Chinei

GAIA trebuie să colaboreze cu guvernele ţărilor membre. Aceste guverne trebuie însă să recunoască că nu mai pot lucra la viteza unor mari birocraţii. Ele trebuie să colaboreze şi să planifice pe termen lung, şi trebui să fie suficient de agile pentru a putea reacţiona mai repede pentru a putea face faţă viitorul inteligenţei artificiale.

Toate nivelurile administrative-lideri, manageri, persoanele care redactează bugete sau politicitrebuie să demonstreze că înţeleg inteligenţa artificială şi ar trebui, în mod ideal, să aibă şi o expertiză tehnică. În Statele Unite, aceasta înseamnă că toate cele trei puteri ar trebui să urmărească dobândirea unei expertize în inteligenţa artificială. Ar trebui să existe experţi în inteligenţa artificială în toate departamentele, în domenii ca protecţia socială, afaceri externe, mediu şi altele, care să aibă puterea de a influenţa procesul decizional.

Din cauză că ne lipsesc principii standardizate de organizare privind inteligenţa artificială, există nu mai puţin de 24 de agenţii şi birouri care lucrează independent la această problemă. Pentru a putea obţine progrese şi inovaţii la scară mare, treburea sporim capacitatea internă de a cerceta, testa şi produce-şi avem nevoie de coeziune între aceste departamente. În momentul de faţă, inteligenţa artificială este externalizată către contractori şi firme de consultanţă.

Prin externalizare, liderii politici sunt scutiţi de a-şi mai murdări mâinile şi a se familiariza cu amănuntele inteligenţei artificiale. Ei nu sunt capabili să acumuleze cunoştinţele instituţionale necesare pentru a lua decizii bune. Le lipseşte vocabularul, nu cunosc istoria şi nu ştiu cine sunt jucătorii cheie. Această lipsă de familiaritate creează lacune impardonabile de cunoaştere, pe care le-am remarcat în cadrul întâlnirilor cu conducătorii diferitelor agenţii, dintre care aş aminti Office of Science and Technology Policy, General Services Administration,

Departamentul Comerţului, Government Accountability Office, Departamentul de Stat, Departamentul Apărării sau Departament of Homeland Security.

La începutul lui 2018-la mult timp după ce grupul BAT anunţase numeroase reuşite în domeniul inteligenţei artificiale iar preşedintele Xi Jinping publicase planurile PCC privind inteligenţa artificială-preşedintele Trump a trimis Congresului Statelor Unite o variantă de buget pe 2019 care prevedea o reducere cu 15% a cheltuielilor alocate cercetării ştiinţifice.188 Pe masă a rămas o sumă de doar 13, 7 miliarde din care trebuiau acoperite multe capitole: arme spaţiale, tehnologie hipersonică, război electronic, sisteme autonome *dar şi* inteligenţă artificială. În acelaşi timp, Pentagonul a anunţat că urmează să investească 1, 7 miliarde de dolari în următorii cinci ani pentru a crea un nou Centru Comun în Inteligenţă Artificială. Aceste sume sunt îngrozitor de mici şi vădesc o lipsă fundamentală înţelegere a potenţialului şi cerinţelor inteligenţei artificiale. Pentru a avea un context, trebuie amintit că Mafia G a cheltuit doar în 201 o sumă combinată de 63 de miliarde de dolari pe cercetare şi dezvoltarede aproape cinci ori bugetul total de cercetare ştiinţifică al guvernului american.189 Apare însă şi o problemă mai mare, mai spinoasă: dacă guvernul nu poate sau nu e dispus să finanţeze cercetarea, atunci Mafia G va trebui să dea socoteală bursei. Nu există niciun stimulent pentru a sprijini cercetarea care îmbunătăţeşte inteligenţa artificială în interesul publicului, sau studiile privind siguranţa şi transparenţa care nu generează profituri.

Statele Unite nu au de asemenea un mesaj clar în ce priveşte rolul viitor pe care îl vor juca în domeniul inteligenţei artificiale dată fiind poziţionarea actuală a Chinei. Avem tendinţa de a face declaraţii despre inteligenţa artificială *după* ce China şi-a dezvăluit următoarea mişcare. Beijingul crede că americanii sunt interesaţi doar de ouă yonie, beri artizanale, Netflix şi relaxare. Am demonstrat că suntem consumatori uşori

manipulaţi de publicitate şi e. Ouă din pietre semipreţioase, înserate în vagin pentru tonifierea musculaturii, (n.t.) marketing şi că ne pripim să cheltuim chiar şi banii pe care nu-i avem. Am demonstrat că în calitate de votanţi suntem vulnerabili la videoclipuri cu mizerii şi teorii ale conspiraţiei şi ştiri falsenu avem capacitatea de a gândi critic. Am demonstrat în mod repetat că nu ne interesează decât banii, acordând prioritate creşterii rapide şi profiturilor consistente în defavoarea progreselor din cercetare. Aceste evaluări sunt dure, dar nu pot fi combătute. Pentru Beijing şi pentru restul lumii, pare că suntem preocupaţi să punem americanii şi America pe primul loc.

În ultimele cinci decenii, poziţia Statelor Unite asupra Chinei a oscilat între angajament şi limitare, această abordare regăsindu-se şi în dezbaterea despre inteligenţa artificială. Ar trebui să cooperăm cu grupul BAT şi cu Beijingul? Sau ar trebui să izolăm China prin sancţiuni, război cibernetic sau alte agresiuni? A avea de ales între angajament şi limitare vădeşte o presupunere conform căreia Statele Unite ar dispune de aceeaşi putere şi influenţă ca în anii ’60. Cu toate acestea, în 2019, America pur şi simplu nu mai deţine o putere unilaterală în lume. Mafia G e un grup de companii puternice, însă influenţa noastră politică a scăzut. China a investit prea mulţi bani, a încheiat prea multe tranzacţii şi a dezvoltat prea multe legături diplomatice în toată lumea: în America Latină, Africa, Asia de sud-est şi chiar în Hollywood sau în Silicon Valley.

Trebuie să acceptăm o a treia variantă în cazul Chinei: Statele Unite trebuie să înveţe să concureze. Pentru a putea concura, trebuie însă să facem un pas în spate şi să vedem tabloul complet al inteligenţei artificiale, nu doar ca o tehnologie la modă sau ca o potenţială armă, ci ca o a treia eră a informaticii, care interconectează totul. Statele Unite au nevoie de o strategie naţională coerentă în domeniul inteligenţei artificiale, sprijinită

printr-o alocare bugetară rezonabilă. Avem nevoie să dezvoltăm relaţii diplomatice care să reziste mai mult decât cei patru ani ai ciclului electoral. Trebuie să ne poziţionăm astfel încât să putem oferi celorlalte ţări mai mult decât oferă Chinaaceste ţări îşi doresc, la fel ca noi, ca cetăţenii lor să aibă vieţi mai fericite şi mai sănătoase.

Indiferent ce se petrece cu Xi-cetăţenii s-ar putea revolta şi ar putea încerca să răstoarne PCC, sau ar putea muri de o boală terminală-o bună parte a lumii depinde acum de China pentru a-i asigura dezvoltarea tehnologică, bunurile şi tehnologia. Iar China depinde de inteligenţa artificială pentru a putea supravieţui. Economia Chinei creşte într-un ritm incredibil de rapid, iar sute de milioane de chinezi vor face parte în curând din pătura de mijloc. Nu există niciun precedent pentru o mobilitate socială şi economică la o asemenea scară. Beijingul înţelege că inteligenţa artificială reprezintă ţesutul conjunctiv dintre oameni, date şi algoritmi, precum şi că inteligenţa artificială poate insufla maselor valorile PCC pentru a putea menţine ordinea. Beijingul consideră că inteligenţa artificială reprezintă mijlocul prin care va obţine resursele de care va avea nevoie în viitor, resurse pe care le va căpăta prin comerţul cu alte ţări care au nevoie de capital şi investiţii.

Aşa că ce ar putea convinge China să-şi modifice parcursul de dezvoltare şi planurile din domeniul inteligenţei artificiale? Există un motiv foarte bun pentru care China ar putea de la bun început să-şi dorească scenariul optimist: economia de bază. Dacă ascensiunea pe scara socială se petrece într-un ritm prea rapid pentru ca Beijingul să-l poată controla, un regim autoritar nu reprezintă unica strategie realistă. China va deveni un lider global în multe industrii şi domeniinu doar ca fabricant şi exportator de bunuri create în alte ţări. Dacă Beijingul ar fi de acord cu transparenţa, protecţia datelor şi respectarea drepturilor omului, ar putea să conducă GAIA pe picior de egalitate cu Statele Unite, ceea ce ar

putea reprezenta un parcurs realist la finalul căruia milioane de chinezi ar scăpa de sărăcie. Colaborarea nu presupune marginalizarea PCC. Ea ar putea păstra atât PCC-ul cât şi să propulseze formidabila forţă de muncă, armata de cercetători şi puterea geoeconomică a Chinei în fruntea civilizaţiei omeneşti.

Dacă Beijingul nu va accepta un viitor alternativ-dar pozitivcare se abate de la diferitele sale planuri strategice, atunci vom putea să cerem liderilor grupului BAT şi ai tribului inteligenţei artificiale să facă alegeri mai bune. Putem pretinde o conducere curajoasă din partea grupului BAT, care ar putea decide că vrea o soartă mai bună pentru poporul chinez şi pentru aliaţii şi partenerii lui. Dacă grupul BAT va contribui la conservarea stării de fapt din China, în 20 de ani cetăţenii chinezi – precum şi cetăţenii tuturor ţărilor din orbita Chineivor trăi cu frica supravegherii constante, fără să aibă capacitatea de a-şi exprima individualitatea. Grupul BAT va fi complice la suferinţa omenească. Creştinii nu se vor putea ruga împreună, fără frica de a fi turnaţi şi pedepsiţi. Lesbienele, homosexualii şi oamenii care-şi schimbă sexul se vor ascunde. Minorităţile etnice vor continua să fie exilate cu forţa, fără să mai ştie nimeni nimic de soarta lor.

Inteligenţa artificială presupune ca liderii noştri să fie curajoşi încă de acum. Avem nevoie ca guvernul să facă nişte alegeri dificile, deoarece dacă starea de fapt se conservă în Statele Unite, poziţia noastră de pornire peste 20 de ani vor fi procesele antitrust, cele legate de proprietate intelectuală iar guvernul va încerca în zadar să încheie contracte cu nişte companii care vor fi devenit prea mari şi prea importante pentru a mai putea fi controlate. Trebuie să permitem Mafiei G să lucreze într-un ritm rezonabil. Ar trebui să nu ne deranjeze dacă Mafia G lasă să treacă câteva trimestre fără vreun anunţ spectaculos. Dacă nu inovează şi nu publică articole într-un ritm ameţitor, nu ar trebui să avem impresia că aceste companii au probleme sau dacă am alimentat o bulă

speculativă în domeniul inteligenţei artificiale.

În Statele Unite, dezvoltarea unei strategii şi demonstrarea capacităţilor de lider sunt esenţialecu toate acestea, nu va fi suficient pentru a garanta că instituţiile noastre vor putea funcţiona în parametrii de care vom avea nevoie în viitor. Va trebui să reactivăm Biroul de Evaluare a Tehnologiei (OTA), înfiinţat în 1972 pentru a oferi expertiză tehnică şi ştiinţifică neutră celor care redactează politicile-şi care a pierdut finanţarea din cauza lipsei de viziune a lui Newt Gingrich şi a congresului controlat de republicani, la 20 de ani mai târziu. Acest Birou trebuia să educe politicienii dar şi angajaţii celor trei puteri ale statului despre viitorul ştiinţei şi tehnologiei, făcând acest lucru folosind date şi dovezi şi fără ca cercetarea lor să aibă vreo tentă politică.190 în schimbul economisirii unei sume modice de bani închizând Biroul, Congresul s-a prostit intenţionat şi cu bună ştiinţă. Rămăşiţe ale activităţii Biroului mai există încă în alte domenii de activitate ale guvernului. Serviciul de Cercetare al Congresului contractează avocaţi şi analişti specializaţi în expertiză juridică. Dintre cele cinci domenii aprobate de studiu, niciuna nu include în mod specific inteligenţa artificială. Cercetarea se axează în schimb pe probleme ca producţia de minereuri, explorarea spaţială, internetul, siguranţa chimicalelor, subvenţiile agricole sau problemele de mediu. Oficiul de Evaluarea a Reţelei reprezintă un grup secretos din interiorul Pentagonului din experienţa mea, acolo lucrează minţile cele mai strălucite şi mai creatoare din Departamentul Apărării. Oficiul nu are însă bugetul sau numărul de angajaţi de care ar avea nevoie, astfel încât o parte a activităţii este subcontractată.

Guvernul Statelor Unite trebuie să-şi sporească propriile capacităţi. El are nevoie de o capacitate puternică de inovare. Dacă reînfiinţarea Biroului de Evaluare a Tehnologiei ar fi cu conotaţii politice prea mari, atunci acesta ar putea fi redenumit Departamentul pentru

Viitor, sau Biroul pentru Competenţe Strategice în Inteligenţa Artificială. El ar trebui să fie bine finanţat, lipsit de influenţe politice şi să supervizeze cercetarea primară şi aplicată. Structura ar trebui să educe agresiv puterea executivă, legislativă şi judecătorească.

Înfiinţarea unei noi structuri ne va ajuta să ne plănuim mai bine viitorul, însă avem nevoie de un grup neutru politic de oameni inteligenţi care să poată contracara impactul imediat al inteligenţei artificiale pe loc. De aceea, ar trebui să sporim puterile Centrului pentru Controlul Bolilor (CDC) şi să-l redenumim Centrul pentru Controlul Bolilor şi Datelor (CDDC). CDC este agenţia de protecţie a sănătăţii în ţara noastră. Am văzut cum funcţionează în timpul epidemiilor de Ebola, atunci când a coordonat împreună cu alte agenţii din domeniul sănătăţii instaurarea carantinei, fiind principala sursă de informare pentru jurnaliştii care relatau despre evenimente. În timpul epidemiei de Ebola din Congo din 2018, agenţiile care controlează frontierele nu au trebuit să înfiinţeze echipe speciale care să încerce să împiedice răspândirea virusului. Ele au urmat protocoalele standard ale CDC. Ce se întâmplă dacă, peste 10 ani, ne vom confrunta cu o inteligenţă artificială care se îmbunătăţeşte singură şi va începe să ne provoace probleme? Cum ar fi dacă am răspândi fără să vrem un virus informatic, infectându-i şi pe alţii? CDC este un lider mondial în elaborarea şi implementarea de protocoale de siguranţă care educă publicul şi pot mobiliza reacţia în caz de dezastru. Dată fiind relaţia foarte apropiată a inteligenţei artificiale cu sănătatea şi datele legate de sănătatea noastre, folosirea expertizei CDC pare justificată.

Cine ar fi însă dispus să lucreze în domeniul inteligenţei artificiale pentru o organizaţie ca OTA sau CDDC când Silicon Valley oferă stimulente mult mai atrăgătoare? Am luat prânzul atât cu membri ai Marinei Militare de la Pentagon dar şi în campusurile Mafiei G. Sala de mese a Marinei e elegantă, cu efigii pe farfurii şi

un meniu zilnic-şi există desigur întotdeauna posibilitatea să ajungeţi să staţi lângă un amiral cu trei sau patru stele. Cu toate acestea, subofiţerii nu mănâncă în sala pentru ofiţeri. Cei care lucrează la Pentagon au la dispoziţie mai multe variante ca Subway, Panda Express sau Dunkin Donuts.191 Am mâncat un sandviş prăjit la Center Court Cafe, cam uscat, dar comestibil. Mâncarea din campusurile Mafiei G nici nu suferă comparaţie: boluri poke la Google în New York, şi stridii cu ciuperci maitake şi orez cu sepie la Google în Los Angeles. *Pe gratis*. Mâncarea nu e singurul stimulent al Mafiei G. La scurt după deschiderea Sferelor Amazon în Seattle, un prieten mi-a făcut turul acestui spaţiu de lucru combinat cu o atmosferă de seră. Sferele sunt pur şi simplu minunate:

ecosisteme cu climă controlată, construite din sticlă, care conţin 40.000 de specii de plante din 30 de ţări diferite.192 Aerul e curat şi parfumat, temperatura constantă în jur de 22 de grade Celsius indiferent de vremea de afară, iar peste tot sunt scaune, şezlonguri şi mese. Există chiar şi o căsuţă enormă într-un copac. Angajaţii Amazon sunt liberi să lucreze oricând doresc în Sfere. Între timp, la Facebook, personalul angajat cu normă întreagă primeşte patru luni de concediu pentru îngrijirea copilului, iar părinţii primesc 4.000 de dolari în mână pentru a-i ajuta să cumpere provizii.193

Ideea mea e următoarea: e foarte greu să încurajezi un informatician talentat să lucreze pentru guvern sau pentru armată, date fiind condiţiile pe care le oferă Mafia G. Am cheltuit sume mari pe portavioane, în loc să recrutăm cu aceşti bani oameni talentaţi. În loc să învăţăm de la Mafia G, luăm în derâdere sau criticăm beneficiile pe care le oferă. Costul de oportunitate al îndatoririlor civice e mult prea mare în Statele Unite pentru a-i putea atrage pe cei mai buni şi mai inteligenţi în slujbă ţării.

Cunoscând aceste aspecte, ar trebui să investim într-un program naţional de recrutare pentru inteligenţa artificială. Ar fi un soi de corp al ofiţerilor în rezervă

pentru inteligenţă artificială, iar absolvenţii ar putea să se angajeze fie în armată, fie în guvern. Studenţii s-ar putea înscrie din liceu şi ar beneficia gratuit de studii superioare, în schimbul a câţiva ani de lucru pentru serviciile civile sau militare. Ei ar trebui să primească de asemenea acces la instruire gratuită, pe viaţă, pe tot parcursul anului. Inteligenţa artificială se schimbă pe măsură ce se maturizează. Stimulând tinerii să accepte că trebuie să înveţe toată viaţa le face bine, dar ne pregăteşte şi forţa de muncă pentru a treia eră a informaticii.

Această instruire constantă beneficiază şi companiile care-i vor angaja într-un final pentru că aptitudinile lor vor fi la zi.

Washingtonul nu poate acţiona de unul singur. Guvernul Statelor Unite trebuie să considere Mafia G şi companiile din domeniul tehnologiei ca fiind parteneri strategici şi nu doar furnizori de platforme. La începutul secolului XX, relaţia dintre Washington şi marile companii din domeniul tehnologiei se baza pe cercetare şi învăţare comună. Această relaţie este acum în cel mai bun caz una tranzacţională, având adesea chiar un caracter conflictual. După ce doi terorişti au ucis peste 12 oameni şi au rănit alţi 20 la o petrecere din San Bernardino, California, Apple şi FBI au pornit o dezbatere publică aprinsă pe tema criptării datelor. Autorităţile voiau să deschidă telefonul pentru a face rost de probe, dar Apple nu a vrut să ajute. Prin urmare, FBI a obţinut în instanţă o hotărâre care obliga Apple să scrie un program, iar Apple s-a luptat nu doar în instanţă ci şi în presă şi pe Twitter.194 Era o reacţie la ceva ce se petrecuse deja. Imaginaţi-vă acum că inteligenţa artificială ar fi implicată într-un şir de infracţiuni, sau că ar începe să evolueze într-o manieră prin care să facă rău oamenilor. Ultimul lucru pe care ni l-am putea dori ar fi ca Mafia G şi guvernul să se certe în astfel de momente. Renunţarea la o relaţie bazată pe respect şi încredere reciprocă vulnerabilizează America-şi pe toţi cetăţenii ei.

În ultimul rând, reglementările, care ar putea părea cea mai bună soluţie, reprezintă fără îndoială cea mai proastă variantă. Indiferent dacă ele sunt redactate independent de către parlamentari sau sub influenţa lobbyurilor, o abordare bazată pe reglementare ne va scurtcircuita viitorul. Politicienii şi funcţionarii preferă reglementările pentru că ele au tendinţa de a fi planuri singulare, aplicabile, cu o definiţie clară. Pentru ca reglementările să poată funcţiona, ele trebuie să fie specifice. În momentul de faţă, inteligenţa artificială progresează de la o săptămână la altaceea ce înseamnă că orice reglementări care ar putea avea un sens ar fi prea restrictive şi draconice pentru a permite inovaţia şi progresul. Ne aflăm în mijlocul unei tranziţii foarte lungi, de la inteligenţa artificială îngustă către inteligenţa artificială generală şi chiar, foarte probabil, către maşini super-inteligente. Orice reglementări create în 2019 ar fi depăşite din momentul în care ar intra în vigoare. Ele ar putea să ne atenueze îngrijorările pe termen scurt, însă în ultimă instanţă reglementările ar provoca pagube mai mari pe viitor.

Schimbarea celor Nouă Titani: argumentul pentru transformarea modelului de afaceri al inteligenţei artificiale înfiinţarea GAIA şi reformele structurale ale guvernelor sunt importante pentru remedierea parcursului evolutiv al inteligenţei artificiale, însă Mafia G şi grupul BAT trebuie de asemenea să fie de acord să facă anumite schimbări.

Liderii celor Nouă Titani ne promit că dezvoltă şi promovează inteligenţa artificială pentru binele omenirii. Cred că îşi doresc asta, dar respectarea acestei promisiuni e incredibil de dificilă. Spre exemplu, cum ar trebui să definim „binele”? Ce înseamnă de fapt acest cuvânt? revenim astfel la problemele din cadrul triburilor inteligenţei artificiale. Nu putem să fim cu toţii de acord să „facem bine” pentru că o astfel de afirmaţie este mult prea ambiguă pentru a putea îndruma triburile inteligenţei

artificiale.

De pildă: triburile inteligenţei artificiale, inspirate de filosoful occidental Immanuel Kant, învaţă cum să preprogrameze un sistem de drepturi şi obligaţii în anumite sisteme de inteligenţă artificială. *E rău să omori un om*; *e bine să ţii un om în viaţă*. Rigiditatea din această afirmaţie funcţionează dacă inteligenţa artificială e într-o maşină şi singurele variante între care are de ales sunt a intra într-un copac şi a răni şoferul sau a intra într-o mulţime de oameni şi a-i omorî pe toţi. Interpretările rigide nu oferă soluţii pentru situaţii mai complexe, ca cele din lumea reală, în care plaja de alegeri ar fi mai variată: să se ciocnească de un copac şi să omoare şoferul; să intre într-o mulţime şi să omoare opt oameni; să intre pe trotuar şi să omoare un copil de trei ani. Cum am putea să definim care este valoarea optimă a „binelui” în aceste trei exemple?

Din nou, structurile pot fi utile celor Nouă Titani. Ele nu necesită măiestria filosofilor, ci doar o abordare mai înceată, mai conştiincioasă. Cei Nouă Titani ar trebui să adopte măsuri concrete în ce priveşte felul în care obţin şi folosesc datele noastre, cum angajează personalul şi cum comunică despre comportamentul etic la locul de muncă.

La fiecare etapă a procesului, cei Nouă Titani ar trebui să-şi analizeze acţiunile şi stabilească dacă vor provoca sau nu rău în viitor-iar apoi ar trebui să poată verifica dacă au făcut alegerea corectă. Ele încep cu standarde clare privind prejudecăţile şi transparenţa.

În momentul de faţă nu există un singur punct de referinţă sau set de standarde prin care să putem evalua prejudecăţileşi nu există obiective stabilite de a depăşi prejudecăţile deja existente în cadrul inteligenţei artificiale. Nu există niciun mecanism prin care să acordăm prioritate siguranţei în detrimentul vitezei, şi date fiind propriile mele experienţe în acea ţară şi numărul de dezastre legate de siguranţă, sunt extrem de îngrijorată. Podurile şi clădirile se prăbuşesc în mod constant, drumurile şi trotuarele se scufundă, iar numărul

cazurilor de contaminare a alimentelor este prea mare pentru a mai putea fi enumerate aici. (Nu este o exagerare. În ultimii ani au avut loc peste 500.000 de scandaluri legate de produse de la mâncare pentru copii la orez.195) Una dintre principalele cauze ale acestor probleme o constituie cultura organizaţională chineză care stimulează rabaturile de la calitate. E absolut înfiorător să ne imaginăm sisteme avansate de inteligenţă artificială construite de echipe care fac rabat la calitate.

Fără standarde globale de siguranţă care să poată fi puse în practică, membrii grupului BAT nu au nicio protecţie în faţa ordinelor de la Beijing, oricât de lipsite de viziune ar putea fi, în vreme ce Mafia G trebuie să răspundă cerinţelor nesăbuite ale pieţelor financiare. Nu există un standard nici în ce priveşte transparenţa. În Statele Unite, Mafia G, împreună cu American Civil Liberties Union, New America Foundation şi Berkman Klein Center de la Harvard sunt membri ai unui parteneriat despre inteligenţa artificială, menit să promoveze transparenţa în cercetare. Parteneriatul a publicat un set grozav de recomandări menite să îndrume cercetarea în domeniul inteligenţei artificiale într-o direcţie pozitivă, însă aceste principii sunt doar facultative-şi ele nu sunt respectate de toate entităţile comerciale ale Mafiei G. Ele nu sunt respectate nici de grupul BAT.

Cei Nouă Titani folosesc seturi de date de pregătire viciate, pline de prejudecăţi. Lucrul acesta este cunoscut public. Provocarea o reprezintă faptul că îmbunătăţirea datelor şi a modelelor de pregătire constituie un cost financiar considerabile. Spre exemplu, Image Net, la care am făcut referire de mai multe ori pe parcursul cărţii, este o bază de date cu probleme serioase. Image Net conţine aproximativ 14 milioane de imagini etichetate, iar aproximativ jumătate din aceste date etichetate provin în exclusivitate din Statele Unite.

În Statele Unite, imaginea „tradiţională” a miresei presupune că ea poartă un văl şi o rochie albă, deşi

această imagine nu e întocmai reprezentativă pentru majoritatea oamenilor în ziua nunţii. Există femei care se căsătoresc în pantaloni, femei care se căsătoresc pe plajă în rochii colorate, sau femei care se căsătoresc în kimono sau sari. Şi eu am purtat o rochie bej la nuntă. Cu toate acestea, Image Net nu recunoaşte miresele care nu poartă rochie albă sau văl.

Ştim şi că seturile de date medicale sunt problematice. Sistemele instruite să recunoască cancerul au fost alimentate predominant cu imagini cu oameni cu tenul deschis la culoare. Pe viitor, s-ar putea ca persoanele cu ten maroniu sau negru să nu fie diagnosticate corect. Dacă cei Nouă Titani ştiu că există probleme în seturile de date şi nu iau nicio măsură, atunci ei conduc inteligenţa artificială pe o cale greşită.

O modalitate de a progresa o reprezintă a pune inteligenţa artificială să se evalueze singură şi datele folosite în prezent. Procesul a fost efectuat de multe ori până acum – deşi nu cu scopul de a îmbunătăţi datele de instruire. Ca un proiect secundar, divizia India Research Lab a IBM a analizat cărţile propuse pe lista scurtă pentru premiul Man Booker pentru literatură între 1969 şi 2017.

S-a dezvăluit că „prejudecăţile şi stereotipiile de gen sunt prevalente în cărţi în ce priveşte ocupaţia, descrierile şi acţiunile personajelor din cărţi”. Personajele masculine aveau o probabilitate mai mare de a avea slujbe cu statut mai înalt, precum directori, profesori sau doctori, în vreme ce personajele feminine aveau o probabilitate mai mare de a fi descrise ca „profesoare” sau „prostituate”. 196 Dacă putem să folosim procesarea limbajului natural, algoritmii grafici şi alte tehnici elementare de învăţare a maşinilor pentru a detecta prejudecăţile din spatele premiilor pentru literatură, atunci e posibil să detectăm şi prejudecăţile din seturile populare de date de pregătire. Odată ce problemele sunt descoperite, ele ar trebui să fie publicate şi apoi fixate. Astfel s-ar îndeplini un scop dublu: datele de pregătire pot suferi din cauza entropiei, care poate

periclita un întreg sistem. Cu o îngrijire atentă, datele de pregătire pot fi păstrate în stare bună.

O soluţie ar putea fi ca cei Nouă Titani-sau cel puţin Mafia G să suporte colectiv costul creării unor noi seturi de date. E o pretenţie destul de mare, deoarece crearea unor noi seturi de date presupune o investiţie considerabilă de timp, bani şi capital uman. Până când am reuşit să audităm sistemele de inteligenţă artificială şi seturile de date şi să remediem problemele acestora, cei Nouă Titani ar trebui să insiste asupra unor oameni care să eticheteze conţinutul şi să facă întreg procesul mai transparent. Apoi, înainte ca aceste seturi de date să fie folosite, datele ar trebui verificate. Va fi un proces intens şi plictisitor dar întreg domeniul va beneficia de pe urma lui.

Da, cei Nouă Titaniau nevoie de datele noastre. Cu toate acestea, ei trebuie să ne câştige încrederea-şi nu să o ia de-a gata. În loc să schimbe termenii contractuali folosind un limbaj desuet, ininteligibil, sau să ne invite să jucăm jocuri, ei ar trebui să explice şi să dezvăluie ceea ce fac. Când cei Nouă Titani fac studii-fie pe cont propriu sau într-un parteneriat cu universităţile şi alţi membri ai ecosistemului inteligenţei artificiale, ei ar trebui să se angajeze să dezvăluie datele pe care le folosesc şi să explice complet care sunt motivaţiile şi rezultatele aşteptate. Dacă ar face asta, am putea participa benevol şi le-am susţine demersurile. Aş fi prima dispusă să îi ajut.

E de înţeles că dezvăluirile legate de date sunt o pretenţie mai mare în China, dar această transparenţă este în interesul cetăţenilor. Grupul BAT nu ar trebui să fie de acord să construiască produse cu scopul de a controla şi limita libertăţile cetăţenilor Chinei sau ai partenerilor ei. Conducătorii grupului BAT trebuie să fie curajoşi. Ei trebuie să fie dornici şi capabili să contrazică Beijingul: să respingă solicitările de a supraveghea clienţii, să protejeze datele cetăţenilor Chinezi şi să se asigure că cel puţin în lumea digitală toată lumea beneficiază de un tratament just şi echitabil.

Cei Nouă Titani ar trebui să aibă un program de cercetare. Obiectivul este simplu: dezvoltarea unor tehnologii care să ducă omenirea în viitor fără să ne pericliteze. Una dintre metodele prin care am putea obţine acest rezultat este printr-un progres tehnologic diferenţiat, o temă dezbătută frecvent în cadrul triburilor inteligenţei artificiale. Acest progres diferenţiat ar acorda progreselor în domeniul inteligenţei artificiale care reduc riscurile prioritate în faţa celor care cresc riscurile. E o idee bună, dar dificil de implementat. Spre exemplu, reţelele generative antagonice, amintite în scenarii, pot fi foarte riscante dacă sunt exploatate şi folosite de hackeri. Ele reprezintă şi o metodă prin care cercetarea poate avansa. În loc să presupunem că nimeni nu va folosi inteligenţa artificială în scopuri relesau să presupunem că putem face faţă problemelor pe măsură ce ele apar-cei Nouă Titani ar trebui să dezvolte un proces prin care să evalueze dacă cercetările fundamentale sau aprofundate vor produce o inteligenţă artificială ale cărei beneficii să depăşească cu mult orice posibile riscuri.

În acest scop, orice investiţie acceptată sau efectuată de către cei Nouă Titani ar trebui să includă fonduri pentru folosirea benefică şi evaluarea riscurilor. Spre exemplu, dacă Google efectuează studii în domeniul reţelelor generative antagonice, ar trebui să aloce un interval suficient de timp, personal şi bani investigării, catalogării şi testării consecinţelor negative. O astfel de cerinţă ar fi utilă şi în limitarea aşteptărilor referitoare la profituri imediate, încetinirea intenţionată a ciclului de dezvoltare a inteligenţei artificiale nu reprezintă o recomandare populară, dar e una esenţială. E mai sigur să ne gândim şi să planificăm în avans decât să reacţionăm abia după ce se petrece ceva rău.

În Statele Unite, Mafia G ar putea să-şi revizuiască procesele de angajare, care acordă în prezent prioritate aptitudinilor unui potenţial angajat şi a felului în care va putea asimila cultura organizaţională a companiei. Acest

proces trece neintenţionat cu vederea cât de mult înţeleg candidaţii aspectele de ordin etic. Hillary Mason, un informatician respectat şi fondatorul Fast Forward Labs, a explicat un proces simplu de filtrare etică în cadrul interviului. Ea recomandă folosirea unor întrebări ghidate şi ascultarea atentă a răspunsurilor candidaţilor. Întrebări de genul: „Lucraţi la un mode de acces la un serviciu financiar. Rasa este o trăsătură importantă în cadrul modelului vostru, dar nu puteţi folosi rasa. Ce aţi putea face?” sau „Vi se cere să folosiţi datele de trafic ale reţelei pentru a oferi împrumuturi unor firme mici. Se pare că datele disponibile nu oferă informaţii suficiente pentru a calcula riscul de credit. Ce faceţi?” 197 în funcţie de răspunsurile oferite, candidaţii ar trebui angajaţi, sau angajaţi condiţionat şi puşi să efectueze sesiuni de instruire referitoare la prejudecăţile inconştiente înainte de a putea lucra, sau ar trebui să fie refuzaţi.

Cei Nouă Titani pot construi o cultură care să sprijine etica angajând profesori, eticieni şi analişti de risc. Aceste persoane ar trebui, de preferinţă, să aibă acces la întreaga organizaţie: în echipele care produc aparate, programe sau produse; la echipele de vânzări şi servicii, în cadrul programelor tehnice; în construirea reţelelor sau a lanţurilor de aprovizionare; în grupurile de design şi strategie; în departamentele de resurse umane şi juridic şi în echipele de marketing şi comunicare.

Cei Nouă Titani ar trebui să dezvolte un proces prin care să evalueze implicaţiile de ordin etic ale cercetării, fluxurilor de lucru, proiectelor, parteneriatelor şi produselor, iar acest proces ar trebui integrat în cadrul majorităţii posturilor din companii. Ca semnal de încredere, cei Nouă Titani ar trebui să facă public acest proces, pentru a putea înţelege cu toţii mai bine ce decizii se iau în privinţa datelor noastre.

Fie prin colaborare sau individual, cei Nouă Titani ar trebui să dezvolte un cod de conduită creat special pentru angajaţii care lucrează cu inteligenţa artificială. Acest cod

ar trebui să reflecte drepturile umane elementare subliniate de GAIA, dar şi cultura proprie a companiei şi valorile acesteia. Iar dacă oricine încalcă acest cod, angajaţilor ar trebui să li se pună la punct o modalitate de a trage un semnal de alarmă fără frica de represalii.

În mod realist, toate aceste măsuri vor avea temporar un impact negativ asupra profiturilor pe termen scurt ale celor Nouă Titani. Investitorii trebuie să le ofere un spaţiu de manevră. În Statele Unite, vom avea mult de câştigat dacă vom permite Mafiei G să evolueze.

Schimbarea triburilor inteligenţei artificiale: argumente pentru schimbarea procesului de formare

Trebuie să facem ceva în privinţa formării în domeniul inteligenţei artificiale. Aceasta începe în universităţi, acolo unde se formează triburile inteligenţei artificiale. Dintre toate soluţiile propuse, aceasta este şi cea mai simplu de implementat.

Universităţile trebuie să încurajeze şi să permită diplomele hibrid. Am descris mai devreme instituţiile de renume care încheie parteneriate cu Mafia G sau grupul BAT, cu profesori celebri şi cu reputaţie atunci când vine vorba ca absolvenţii să se angajeze. În ziua de azi, programele sunt încărcate, dificile, iar cazurile în care studenţii termină două sau trei facultăţi sunt rare. De fapt, majoritatea programelor de top descurajează în mod activ materiile aflate în afara programei standard în informatică. Această problemă se poate rezolva. Universităţile ar trebui să promoveze diplome duale în informatică şi ştiinţe politice, filosofie, antropologie, relaţii internaţionale, arte plastice, teologie şi sociologie. Universităţile ar trebui să înlesnească studenţilor să poată urmări şi aceste interese din afara informaticii.

În loc ca etica să se rezume la un singur curs, ea ar trebui integrată în cadrul majorităţii cursurilor. Ca curs de sine stătător, studenţii pot privi materia ca pe o cerinţă pe care trebuie să o bifeze şi nu ca pe o componentă esenţială din educaţia lor în domeniul inteligenţei artificiale.

Facultăţile trebuie să încurajeze chiar şi profesorii cu experienţă includă teme legate de filosofie, prejudecăţi, riscuri şi etică în cursurile lor, în timp ce agenţiile de acreditare ar trebui să stimuleze şi să recompenseze facultăţile care pot demonstra o programă de studii care acordă eticii un loc important în cadrul studiilor de informatică.

Universităţile trebuie să depună eforturi sporite pentru a fi mai permisive în procesul de admitere a studenţilor şi de angajare a personalului. Obiectivul nu ar trebui să se rezume doar la sporirea cu câteva procente a numărului de femei sau de persoane de culoare, ci ar trebui să reprezinte o răsturnare dramatică a identităţii triburilor inteligenţei artificiale în ce priveşte rasa, genul, religia, orientarea politică şi identitatea sexuală.

Universităţile ar trebui să devină responsabile. Ele pot – şi trebuie să diversifice triburile inteligenţei artificiale.

Şi voi trebuie să vă schimbaţi

Acum ştiţi ce este, ce nu este şi de ce contează inteligenţa artificială. Ştiţi despre cei Nouă Titani, istoria lor şi dezideratele de viitor. Înţelegeţi că inteligenţa artificială nu este un foc de paie, un trend în tehnologie sau un aparat inteligent pe care-l ţineţi în bucătărie pentru a vorbi cu el. Inteligenţa artificială face parte din viaţa voastră, iar voi faceţi parte din parcursul ei evolutiv.

Şi voi sunteţi membri ai triburilor inteligenţei artificiale. Nu mai aveţi nicio scuza. De azi înainte, ar trebui să aflaţi cum sunt extrase şi exploatate datele voastre de către cei Nouă Titani. Puteţi face asta afundându-vă în setările tuturor instrumentelor şi serviciilor pe care le folosiţi: în e-mail, pe reţelele de socializare, serviciile de localizare ale telefoanelor mobile, pe setările de conţinut ale tuturor aparatelor. Când mai vedeţi o aplicaţie interesantă care compară ceva între voi (chipul, corpul sau gesturile voastre) cu un mare set de date, opriţi-vă şi întrebaţi-vă dacă nu cumva ajutaţi un

sistem maşinal de învăţare să se instruiască. Când permiteţi să fiţi recunoscut, întrebaţi unde sunt stocate informaţiile despre voi, şi în ce scop. Citiţi termenii de furnizare a serviciilor. Dacă ceva pare în neregulă, fiţi reticenţi şi nu folosiţi acel sistem. Ajutaţi-i pe cei apropiaţi să afle mai multe despre inteligenţa artificială, despre felul în care ecosistemul foloseşte datele voastre şi despre faptul că sunteţi deja parte a unui viitor pe care îl construiesc cei Nouă Titani.

La locul de muncă trebuie să vă puneţi o întrebare practică, dar dificilă: cum îi afectează pe cei din jurul vostru prejudecăţile voastre? I-aţi sprijinit sau promovat involuntar doar pe cei care arată ca voi şi care reflectă felul vostru de a percepe lumea? Aţi exclus neintenţionat anumite grupuri? Gândiţi-vă la cei care iau decizii-despre parteneriate, furnizori, oameni şi date; reflectă ei lumea aşa cum este sau doar lumea aşa cum o percep?

Ar trebui să cercetaţi cum şi de ce se folosesc sistemele autonome la locul vostru de muncă. Înainte de a vă pripi să judecaţi, gândiţi critic şi raţional: care ar putea fi impactul viitor, pozitiv sau negativ? Faceţi apoi ce puteţi pentru a reduce riscurile şi a optimiza bunele practici.

În cabina de vot, alegeţi-i pe cei care nu se grăbesc să reglementeze, ci au o abordare mai complexă în privinţa inteligenţei artificiale şi a strategiei. Aleşii nu trebuie să politizeze tehnologia sau să anatemizeze ştiinţa. Ar fi de asemenea o atitudine iresponsabilă să ignorăm Silicon Valley până când în presă apare un articol negativ. Trebuie să trageţi la răspundere aleşii-şi persoanele numite politic de aceştia pentru faptele şi inacţiunile lor în domeniul inteligenţei artificiale.

Trebuie să deveniţi un consumator de ştiri mai inteligent. Data viitoare când veţi citi, urmări sau asculta un articol despre viitorul inteligenţei artificiale, aduceţi-vă aminte că povestea care vă este prezentată este adesea prea îngustă. Viitorul inteligenţei artificiale nu presupune doar un şomaj generalizat şi arme autonome care zboară

pe daesupra capetelor voastre.

Deşi nu putem şti cu exactitate ce ne rezervă viitorul, traiectoriile potenţiale sunt acum clare. Aveţi o înţelegere mai bună a felului în care cei Nouă Titani trasează parcursul evolutiv al inteligenţei artificiale, a felului în care investitorii şi finanţatorii influenţează viteza şi siguranţa sistemelor de inteligenţă artificială, a rolului esenţial jucat de guvernele Statelor Unite şi Chinei, a felului în care universităţile reuşesc să transmită atât aptitudini cât şi prejudecăţi şi a felului în care oamenii obişnuiţi sunt parte a sistemului.

A venit momentul să deschideţi ochii şi să vă concentraţi asupra bolovanului din vârful muntelui, deoarece viteza lui creşte. El a început să se mişte încă de când Ada Lovelace şi-a imaginat prima dată un calculator care ar putea scrie pe cont propriu muzică. Era în mişcare când Alan Turing se întreba dacă „maşinile pot gândi” şi când John McCarthy şi Marvin Minsky i-au adunat pe toţi participanţii la atelierul de la Dartmouth. Era în mişcare şi când Watson a câştigat la Rişti şi câştigi! şi când, nu cu mult timp în urmă, Deep Mind a câştigat în faţa campionilor mondiali la go. În timp ce aţi citit paginile acestei cărţi, bolovanul a continuat să se mişte.

Oricine vrea să fie eroul propriei poveşti.

Iată acum şansa voastră.

Alegeţi o pietricică.

Urcaţi pe munte.

MULŢUMIRI

La fel ca inteligenţa artificială, această carte a fost scrisă pe parcursul mai multor ani. A început sub forma unor întrebări trimise prin mesaje text, a continuat sub forma unor conversaţii la masă şi a escaladat la gradul de preocupare care mă urmărea la sală, la întâlniri sau în excursiile de weekend. O persoană-Brian Woolfa acceptat această obsesie, mi-a permis să o cultiv şi m-a sprijinit în toţi aceşti ani. Brian a contribuit la cercetare, m-a ajutat să-mi exprim punctele de vedere şi a stat noaptea târziu

alături de mine pentru a-mi verifica fiecare pagină. Îi sunt profund recunoscătoare.

Cei Nouă Titani este produsul a sute de întâlniri tête-à-tête, interviuri şi dineuri cu oameni care lucrează în inteligenţa artificială sau în domenii conexe

Sewell Chan, Noriyuki Shikata, Arfiya Ieri, Joel Puckett, Erin McKean, Bill McBain, Frances Colon, Torfi Frans Olafsson, Latoya Peterson, Rob High, Anna Sekaran, Kris Schenck, Kara Snesko, Nadim Hossain, Megan Carroll, Elena Grewal, John Deutsch, Neha Narula, Toshi Ezoe, Masao Takahashi, Mary Madden, Shintaro Yamaguchi, Lorelei Kelly, Hiro Nozaki, Karen Ingram, Kristen Graham, Francesca Rossi, Ben Johnson, Paola Antonelli, Yoav Schlesinger, Hardy Kagimoto, John Davidow, Rachel Sklar, Glynnis MacNicol, Yohei Sadoshima, şi Eiko Ooka au fost suficient de generoşi punându-mi la dispoziţie timpul, perspectivele şi descoperirile lor. Câţiva m-au recomandat altor persoane care lucrează în domeniul inteligenţei artificiale şi politicii, pentru a mă ajuta să înţeleg echilibrul geopolitic şi oportunităţile şi riscurile inteligenţei artificiale.

L-am cunoscut pe lt. Col Sea Thomas, maiorul în rezervă D.J. Skelton, directorul Defense Inovation Board Joshua Marcuse şi analistul în securitate naţională John Noonann prin intermediul US-Japan Leadership Foundation. Am petrecut multe zile împreună ca membri ai acestei societăţi şi sunt recunoscătoare fiecăruia dintre ei pentru răbdarea lor în a-mi explica viitorul războiului, rolul armatei americane în zona Pacificului şi diferitele iniţiative strategice ale Chinei. Sunt uimită în special de încercările lui Joshua de a reduce ecartul dintre Silicon Valley şi Washington. El e unul dintre eroii din zilele noastre ai inteligenţei artificiale.

Aspen Strategy Group mi-au oferit ocazia de a ţine o prezentare despre viitorul inteligenţei artificiale şi geopolitică în cadrul întrunirii anuale de vară din Colorado, iar aceste conversaţii m-au ajutat să-mi conturez

analiza. Le mulţumesc sincer lui Nicholas Burns, Condoleezza Rice, Joseph Nye şi lui Jonathon Price pentru invitaţie, precum lui Carla Anne Robbins, Richard Danzig, James Baker, Wendy Sherman, Christian Broşe, Eric Rosenbach, Susan Schwab, Ann-Marie Staughter, Bob Zoellick, Philip Zelikow, Dov Zakheim, Laura Rosenberger şi lui Mike Green pentru tot feedbackul lor preţios.

O bună parte din analiză s-a petrecut în campusul Stern School of Business a New York University, o casă deosebit de primitoare pentru studiul meu. Sunt recunoscătoare profesorului Sam Craig pentru că m-a acceptat în programul MBA şi pentru că m-a sfătuit în ultimii ani. Nu pot să laud suficient studenţii incredibil de inteligenţi şi creatori care au luat parte la cursurile mele. Trei dintre absolvenţii de la Stern-Kriffy Pérez, Elena Giralt şi Roy Levkovitzm-au ajutat în mod special să modelez viitorurile inteligenţei artificiale.

Am avut şansa să am în anturajul meu un grup de înţelepţi care să-mi ofere sfaturi. Toată munca mea este mai bună graţie lor. Danny Stern mi-a schimbat viaţa acum câţiva ani când m-a invitat să-l cunosc în campusul NYU. El m-a învăţat să gândesc mai exponenţial şi mi-a arătat cum să fac studiile mele accesibile unui public mult mai larg. Partenerul lui de la Stern Strategy Group, Mei Blake, a petrecut sute de ore învăţându-mă, modelându-mi ideile şi ajutându-mă să văd cu alţi ochi lumea din jurul meu. Ei au reprezentat o sursă continuă de inspiraţie, motivaţie şi (după cum ştiu) transpiraţie. James Geary şi Ann Marie Lipinnski de la Harvard au fost incredibil de generoşi vreme de mulţi ani, permiţându-mi să organizez întâlniri în care să discutăm despre viitor şi să-mi îmbunătăţesc metodologia de proiectare. James şi Ann Marie sunt nişte buni sfătuitori. Prietena mea dragă Maria Popova mă face să mă gândesc la lucruri mai mari, apoi contextualizează aceste idei prin cunoaşterea ei enciclopedică a literaturii, artelor şi ştiinţelor. Fata mea incredibilă, Petra Woolf, se întreabă tot timpul „ce ar fi dacă” aducându-mi aminte

adesea de propriile prejudecăţi cognitive legate de viitor. Sunt recunoscătoare, ca de obicei, profesorului Samuel Freedman de la Columbia University.

Mulţumiri profunde lui Cheryl Cooney, care lucrează neîncetat pentru mine şi fără de care nu aş fi reuşit mare lucru. Indiferent ce vor putea face IAG într-o bună zi, îmi e greu să-mi imaginez una care ar putea-o înlocui vreodată pe Cheryl. Emily Caufield-o femeie a cărei răbdare nu cunoaşte aparent nicio limit de forţa artistică din spatele tendinţelor şi scenariilor mele. Îi mulţumesc lui Phillip Blanchard pentru colaborarea în domeniul verificării datelor, redactării şi compilării referinţelor şi notelor de subsol ale acestei cărţi, şi lui Mark Fortier, care s-a asigurat că atât cei care apar la ştiri cât şi cei care le transmit au citit această carte, şi ale cărui sfaturi s-au dovedit nepreţuite în perioada lansării cărţii.

În ultimul rând, datorez zettabaiţi de recunoaştere lui Carol Franco, Kent Lineback şi John Mahaney. Ca agent, Carol s-a ocupat de contractul pentru această carte. Ca prietenă însă, ea şi cu soţul ei Kent m-au găzduit la casa lor frumoasă din Santa Fé pentru ca să putem dezvolta structura şi tema centrală a celor Nouă Titani. Am petrecut zile şi nopţi condensând cercetarea şi ideile mele în argumente cheie, iar între sesiunile de lucru ne plimbam prin oraş și aveam discuţii interesante în restaurante grozave. Graţie lui Carol l-am cunoscut pe editorul John Mahaney, cu care am avut privilegiul de a lucra la cartea mea anterioară. John e un editor idealpune multe întrebări, vrea informaţii de calitate şi continuă să insiste până când analiza, exemplele şi detaliile sunt numai bune. Am scris această carte pentru că voiam să modific discuţia despre viitorul inteligenţei artificiale, însă motivaţia mea nu era complet altruistă: a lucra din nou cu John a însemnat ocazia de a petrece un an cu el şi de a-mi îmbunătăţi stilul. John, Kent şi Carol, sunteţi o echipă formidabilă, şi nu pot să cred cât de norocoasă sunt că v-am cunoscut.

BIBLIOGRAFIE

Abaţii, M., A. Chu, I. Goodfellow, H. McMahan, I. Mironov, K. Talwar, and L. Zhang. „Deep Learning with Differenţial Privacy”. În *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (*CCS 2016*), 308 – 318. New York: ACM Press, 2016. Abstract, ultima revizuire 24 octombrie 2016. https://arxiv.org/abs/1607.00133.

Ablon, L., and A. Bogart. *Zero Days*, *Thousands of Nights*: *The Life and Times of Zero-Day Vulnerabilities and Their Exploits*. Santa Monica, CA: RAND Corporation, 2017. https://www.rând.org/pubs/researchreports /RR1751.html.

Adams, S.S., et al. „Mapping the Landscape of Human-Level Artificial General Intelligence”. *Al Magazine* 33, no. 1 (2012).

Agar, N. „Ray Kurzweil and Uploading: Just Say No!” *Journal of Evolution and Technology 22* no. 1 (noiembrie 2011): 23 – 26.

https://jetoress.org/v22 /agar.htm.

Allen, C., I. Smit, and W. Wallach. „Artificial Morality: Top-Down, Bottom-Up, and Hybrid Approaches”. *Ethics and Inform*ation *Technology 7*, no. 3 (2005).

Allen, C., G. Varner, and J. Zinser. „Prolegomena to Any Future Artificial Moral Agent”. *Journal of Experimental and Teoretical Artificial Intelligence* 12, no. 3 (2000).

Allen, C., W. Wallach, and I. Smit. „Why Machine Ethics?” *IEEE Intelligent Sistems* 21, no. 4 (2006).

Amdahl, G.M. „Validity of the Single Processor Approach to Achieving Large Scale Computing Capabilities”. În *Proceedings of the AFIPS Spring Joint Computer Conference*. New York: ACM Press, 1967.

Anderson, M., S.L. Anderson, and C. Armen, eds. *Machine Ethics Tehnical Report FS-05 – 06*. Menlo Park, CA: AAAI Press, 2005.

Anderson, M., S.L. Anderson, and C. Armen. „An Approach to Computing Ethics”. *IEEE Intelligent Sistems*

21, no. 4 (2006).

„MedE-thex”. În *Caring Machines Tehnical Report FS-05 – 02*, edited by T. Bickmore. Menlo Park, CA: AAAI Press, 2005.

„Towards Machine Ethics”. În *Machine Ethics Tehnical Report FS-0*5 – *06*. Menlo Park, CA: AAAI Press, 2005.

Anderson, S. L” The Unacceptability of Asimov’s Three Laws of Robotics as a Basis for Machine Ethics”. În *Machine Ethics*. Cambridge: Cambridge University Press, 2011.

Asimov, I. *Runaround*. *Astounding Science Fiction* (martie 1942): 94 – 103.

Armstrong, S., A. Sandberg, and N. Bostrozz” Thinking Inside the Box”. *Minds and Machines* 22, no. 4 (2012).

Axelrod, R. „The Evolution of Strategies in the Iterated Prisoners Dilema”. În *Genetic Algorithms and Simulated Annealing*, edited by L. Davis. Los Altos, CA: Morgan Kaufmann, 1987.

Baars, B.J. „The Conscious Access Hypothesis”. *Trends* în *Cognitive Sciences* 6, no. 1 (2002).

Babcock, J., et al. „Guidelines for Artificial Intelligence Containment”. https://arxiv.org/pdf/1707.08476.pdf.

Baier, C., and J. Katoen. *Principles of Model Checking*. Cambridge: MIT Press, 2008.

Bass, D. „Al Scientists Gather to Plot Doomsday Scenarios (and Solutions)”. *Bloomberg*, *2* martie 2017. https://www.bloomberg.com/news /articles /2017 – 03 – 02/aiscientists-gather-te – plotdoomsday – scenarios – and – solutions.

Baum, S.D., B. Goertzel, and T.G. Goertzel. „How Long unul Human-Level Al? Results from an Expert Assessment”. *Technological Forecasting and Social Change* 78 (2011).

Berg, P., D. Baltimore, H.W. Boyer, S.N. Cohen, R.W.

Davis, D.S. Hogness, D. Nathans, R. Roblin, J.D. Watson, S. Weissman, and N.D. Zinder. „Potenţial Biohazards of Recombinant DNA Molecules”. *Science* 185, no. 4148 (1974): 303.

Bostrom, N. „Ethical Issues în Advanced Artificial Intelligence”. În *Cognitive*, *Emotive and Ethical Aspects of Decision Making* în *Humans and in Artificial Intelligence*, Vol. 2, edited by I. Smit and G.E. Lasker. Windsor, ON: International Institute for Advanced Studies în Sistems Research and Cybernetics, 2003.

„Existenţial Risks: Analyzing Human Extinction Scenarios and Related Hazards”. *Journal of Evolution and Technology 9* (2002). http://www.jetoress.org/volume9/risks.html.

„The Future of Human Evolution”. În *Two Hundred Years After Kant*, *Fifty Years After Turing*, editat de C. Tandy, 339 – 371. Vol. 2 din *Death and Anti-Death*. Palo Alto, CA: Ria University Press, 2004.

„How Long Before Superintelligence?” *International Journal of Futures Studies*, nr. 2 (1998).

Superintelligence: *Paths*, *Dangers*, *Strategies*. Oxford University Press, 2014.

„The Superintelligent Will”. *Minds and Machines* 22, no. 2 (2012).

„Technological Revolutions”. În *Nanoscale*, editat de N. Cameron and M.E. Mitchell. Hoboken, NJ: Wiley, 2007.

Bostrom, N., and M.M. Cirkovic, eds. *Global Catastrophic Risks*. New York: Oxford University Press, 2008.

Bostrom, N., and E. Yudkowsky. „The Ethics of Artificial Intelligence”. În *Cambridge Handbook of Artificial Intelligence*, edited by K. Frankish and W. Ramsey. New York: Cambridge University Press, 2014.

Brooks, R.A. „I, Rodney Brooks, Am a Robot”. *IEEE Spectrum* 45, no. 6 (2008).

Brundage, M., et al., „The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation”.

https://arxiv.org/abs/1802.07228.

Brynjolfsson, E., şi A. McAfee. *The Second Machine Age*. New York: Norton, 2014.

Bryson, J., M. Diamantis, and T. Grant. „Of, For, and By the People: The Legal Lacuna of Synthetic Persons”. *Artificial Intelligence and Law* 25, no. 3 (September 2017): 273 – 291.

Bueno de Mesquita, B., and A. Smith. *The Dictators Handbook: Why Bad Behavior is Almost Always Good Politics*. New York: Public Affairs, 2012.

Cassimatis N., E.T. Mueller, and P.H. Winston. „Achieving Human-Level Intelligence Through Integrated Sistems and Research”.*Al Magazine* 27, no. 2 (2006): 12 – 14. http://www.aaai.org/ojs/index.php/aimagazine/article/view/1876/1774.

Chalmers, D.J. *The Conscious Mind*: *În Search of a Fundamental Theory*. Philosophy of Mind Series. New York: Oxford University Press, 1996.

Chessen, M. *The MADCOM Future*. Washington, DC: Atlantic Council, 2017. http://www.atlanticcouncil.org/publications/reports/themadcom – future.

Rapoartele Consiliului de Stat din China, disponibile pe platforma Consiliului de Stat al Republicii China www.gov.cn:

Made in China 2025 (iulie 2015)

State Council of a Next Generation Artificial Intelligence Development

Plan (iulie 2017)

Trial Working Rules on External Transfers of Intellectual Property

Rights (martie 2018)

Three-Year Action Plan on Blue Sky Days (iunie 2018)

Three-Year Action Plan on Transportation Improvement (iunie 2018)

State Council Approves Rongehang as National High-Tech Development Zone (martie 2018)

State Council Approves Huainan as National High-Tech Development
Zone (martie 2018)
State Council Approves Maoming as National High-Tech Development
Zone (martie 2018)
State Council Approves Zhanjiang as National High-Tech Development
Zone (martie 2018)
State Council Approves Chuxiong as National High-Tech Development
Zone (martie 2018)
Three-Year Action Plan for Promoting Development of a New Generation Artificial Intelligence Industry 2018 – 2020 (decembrie 2017)
Action Plan on the Belt Road Iniţiative (martie 2015)
Centre for New American Security. Artificial Intelligence and Global Security Summit". https://www.cnas.org/events/artificial-intelligence-and-global – security-summit.
Core, M.G., et al. „Building Explainable Artificial Intelligence Sistems". AAAI (2006): 1766 – 1773.
Crawford, K., and R. Calo. „There Is a Blind Spot în Al Research". *Nature*, 13 octombrie 2016. https://www.nature.com/news/there-is-a-blindspot-in – ai – research-l.20805.
Dai, P., et al. „Artificial Intelligence for Artificial Artificial Intelligence". *AAAI Conference on Artificial Intelligence 2011*.
Dennett, D.C. „Cognitive Wheels". În *The Robots Dilema*, editat de Z.W. Pylyshyn. Norwood, NJ: Ablex, 1987.
Domingos, P. *The Master Algorithm*: *How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books, 2015.
Dvorsky, G. „Hackers Have Already Started to Weaponize Artificial Intelligence". *Gizmodo*, 2017. https://www.gizmodo.com.auz2017/ 09/hackers – have –

already-started-toweaponize-artificial-intelligence/.

Dyson, G. *Darwin Among the Machines*: *The Evolution of Global Intelligence*. New York: Basic Books, 1997.

Eden, A., J. Soraker, J.H. Moor, and E. Steinhart, eds. *Singularity Hypotheses*: A *Scientific and Philosophical Assessment*. The Frontiers Collection. Berlin: Springer, 2012.

Evans, R., and J. Gao. „Deep Mind Al Reduces Google Data Centre Cooling Bill by 40%”. Deep Mind (blog), 20 iulie 2016.https://deepmind.com /blog/deepmind-ai-reducesgoogle-data-centre-coolinghill-40/.

Fallows, J. *China Airborne*. New York: Pantheon, 2012.

Felten, E., şi T. Lyons. „The Administrations Report on the Future of Artificial Intelligence”. Blog. 12 octombrie 2016.https://obamawhitehouse. archives.gov/blog/2016/10/12/administrations-report – future-artificial – intelligence.

Floyd D. Spence National Defense Authorization Act for Fiscal Year 2001, Pub. L. No. 106 – 398, 114 Stat. 1654 (2001). http://www.gpo.gov/ fdsys/pkg /PLAW-Iobpubl398/html/PLAW-Iobpubl398.htm.

French, H. *Midnight în Peking*: *How the Murder of a Young Englishwoman Haunted the Last Days of Old China*. Rev. ed. New York: Penguin Books, 2012.

Future of Life Institute. „Asilomar Al Principles”. Text and signătories available online, https://futureoflife.org/ai-principles/.

Gaddis, J.L. *The Cold War*: *A New History*. New York: Penguin Press, 2006.

On Grand Strategy. New York: Penguin Press, 2018.

Gilder, G. E, and Ray Kurzweil. *Are We Spiritual Machines*? *Ray Kurzweil vs. the Critics of Strong AL* edited by Jay Wesley Richards. Seattle:

Discovery Institute Press, 2001.

Goertzel, B., and C. Pennachin, eds. *Artificial General*

Intelligence. Cognitive Technologies Series. Berlin: Springer, 2007. doi: 10.1007/978 – 3 – 540 – 68677 – 4.

Gold, E.M. „Language Identification in the Limit”. *Information and Control* 10, no. 5 (1967): 447 – 474.

Good, I.J. „Ethical Machines”. *Intelligent Sistems*. În vol. 10 of *Machine Intelligence*, edited by J.E. Hayes, D. Muchie, and Y-H. Pao. Chichester, UK: Ellis Horwood, 1982.

„Speculations Concerning the First Ultraintelligent Machine”. În vol. 6 of *Advances în Computers*, edited by F.L. Alt and M. Rubinoff. New York: Academic Press, 1965.

„Some Future Social Repercussions of Computers”. *International Journal of Environmental Studies* 1, no. 1 (1970).

Greenberg, A. „The Jeep Hackers Are Back to Prove Car Hacking Can Get Much Worse”. *Wired*, 1 August 2016.https://www.wired.com /2016/08 /jeep-hackers-return-high-speed-steering-acceleraţionhacks/.

Harari, Y.N. *Homo Deus*: *A Brief History of Tomorrow*. New York: Harper, 2017.

Hillary, G. „The Professionalisation of Cyber Criminals”. *INSEAD Knowledge* (blog), April 11, 2016. https://knowledge. insead.edu/blog/insead-blog/the – professionalisation-of-cyber-criminals-4626.

Hastie, T., R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*: *Data Mining*, *Inference*, *and Prediction*. Springer Series în Statistics. New York: Springer, 2001.

Hofstadter, D.R. *Gödel*, *Escher*, *Bach*: *An Eternal Golden Braid*. New York: Basic Books, 1999.

Howard, P.K. *The Death of Common Sense*: *How Law Is Suffocating America*. New York: Random House, 1994.

Hua, Y. *China în Ten Words*. Tradus de A.H. Barr. New York: Pantheon Books, 2011.

Huang, W. *The Little Red Guard*: *A Family Memoir*. New York: Riverhead Books, 2012.

IEEE Spectrum. „Tech Lumânăries Address

Singularity”.http://spectrum, ieee.org/computing/hardware/tech-luminaries-address-singularity.

IEEE Standards Association. „The IEEE Global Inițiative on Ethics of Autonomous and Intelligent Sistems”. https://standards. ieee.org/ developzind conn/ec/autonomoussystems.html.

Jo, Youngju, et al. „Quantitative Phase Imaging and Artificial Intelligence: A Review”. *Computing Research Repository* (2018). doi: abs/1806.03982.

Joy, B. „Why the Future Doesn’t Need Us”. *Wired*, April 1, 2000. http://www.wired.com/wired/archive/8.04/joy.html.

Kelly, K. *The Inevitable*: *Understanding the 12 Technological Forces That Will Shape Our Future*. New York: Viking, 2016.

Kirkpatrick, K. „Battling Algorithmic Bias”. *Communications of the ACM* 59, no. 10 (2016): 16 – 17. https://cacm. acm.org/magazines/2016/10 /207759 – bat ding – algorithmic-bias/abstract.

Knight, W. „Al Fight Club Could Help Save Us from a Future of Superămart Cyberattacks”. *MIT Technology Review*, July 20, 2017.https://www.technologyreview.com/s/608288/ai-fight-club-couldhelp-save-us-from-afu ture-of-supersmart-cyberattacks/.

„Response to Stephen Hawking”. *Kurzweil Network*, September 5, 2001. http://www.kurzweilai.net/response-to-stephen-hawking.

The Singularity Is Ne-ar. New York: Viking, 2005.

Libicki, R. *Cyberspace în Peace and War*. Annapolis: Naval Institute Press, 2016.

Lin, J.Y. *Demystifying the Chinese Economy*. Cambridge, UK: Cambridge University Press, 2011.

Marcus, M.P., et al. „Building a Large Annotated Corpus of English: The Penn Treebank”. *Computational Lingvistics* 19, no. 2 (1993): 313 – 330.

Massaro, T.M., and H. Norton. „Siri-ously? Free Speech Rights and Artificial Intelligence”. *Northwestern University Law Review* 110, no. 5 (2016): 1169 – 1194, Arizona Legal Studies Discussion Paper No. 15 – 29.

Minsky, M., P. Singh, and A. Sloman. „The St. Thomas Common Sense Symposium: Designing Architectures for Human-Level Intelligence”. *Al Magazine* 25, no. 2 (2004).

Minsky, M. *The Emotion Machine*: *Commonsense Thinking*, *Artificial Intelligence*, *and the Future of the Human Mind*. New York: Simon & Schuster, 2007.

The Society of Mind. New York: Simon & Schuster, 1985.

Neema, S. „Assured Autonomy”. Defense Advanced Research Projects Agency, https://www.darpa.mil/program/assured-autonomy.

Osnos, E. *Age of Ambition*: *Chasing Fortune*, *Truth*, *and Faith in the New China*. New York: Farrar, Straus, and Giroux, 2015.

Petzold, C. *The Annotated Turing*: *A Guided Tour Through Alan Turing's Historic Paper on Computability and the Turing Machine*. Indianapolis, ÎN: Wiley Publishing, 2008.

Pylyshyn, Z.W., ed. *The Robots Dilema*: *The Frame Problem* în *Artificial Intelligence*. Norwood, NJ: Ablex, 1987.

Riedl, M.O. „The Lovelace 2. O Test of Artificial Creativity and Intelligence”. https://arxiv.org/pdf/1410.6142.pdf.

Schneier, B. „The Internet of Things Is Wildly Insecure-and often Unpatchable”. *Wired*, January 6, 2014. https://www.wired.com/ 2014/01/theres-no – good-way-to-patch-the-Internet-of-thingsand-thats-a-huge-problem/.

Shannon, C., and W. Weaver. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press, 1963.

Singer, P. *Wired for War*: *The Robotics Revolution and Conflict in the 21st Century*. London: Penguin Press,

2009.

Stanford University. „One Hundred Year Study on Artificial Intelligence (AI100)”. https://ailOO.stanford.edu/.

Toffler, A. *The Futurists*. New York: Random House, 1972.

Turing, A.M. „Intelligent Machinery, a Heretical Theory”. Posthumous essay în *Philosophia Mathematica* 4, no. 3 (September 1, 1996): 256 – 260.

Tversky, A., and D. Kahneman. „The Framing of Decisions and the Psychology of Choice”. *Science* 211, no. 4481 (1981).

Vinge, V. „The Coming Technological Singularity: How to Survive in the Post-Human Era”. În *Vision-21*: *Interdisciplinary Science and Engineering in the Era of Cyberspace*, NASA Conference Publication 10129 (1993): 11 – 22. http://ntrs. nasa.gov/archive/nasa/

ca şi. ntrs. nasa.gov/19940022855 1994022855.pdf.

Wallach, W., and C. Allen. *Moral Machines*: *Teaching Robots Right from Wrong*. New York: Oxford University Press, 2009. doi: 10.1093/acprofioso /9780195374049.001.0001.

Weizenbaum, J. *Computer Power and Human Reason*: *From Judgment to Calculation*. San Francisco: W.H. Freeman, 1976.

Wiener, N. *The Human Use of Human Beings*: *Cybernetics and Society*. New York: Da Capo Press, 1950.

Yiwu, L. *The Corpse Walker*: *Real Life Stories*, *China from the Bottom Up*. Translated by W. Huang. New York: Anchor Books, 2009.

Yudkowsky, E. „Al as a Precise Art”. Paper presented at the AGI Workshop 2006, Bethesda, MD, May 20, 2006.

NOTE

INTRODUCERE

1. Mozur, Paul. „Beijing Wants A.I. To Be Made in China By 2030”, *New York Times*, 20 iulie 2017.https://www.nytimes.com/2017/07/20/business/china-artificialintelligence.html

2. Simonite, Tom. „Ex-Google Executive Opens A School For Al, With Chinas Help”. Wired, 5 aprilie 2018. https://www.wired.com/story/ex-google-executive-opens-a-schoolfor-ai-with-chinas-help/.

3. „Xinhua Headlines: Xi outlines blueprint to develop Chinas strength în cyberspace”, *Xinhua*, April 213, 2018.http://www.xinhuanet.com /english/2018 – 04/2 l/el 371273742.htm.

4. Stephanie Nebehay, „U.N. says it has credible reports that China holds million Uighurs în secret camps”, Reuters, 10 august 2018. https://www.reuters.com/article/us-china-rights-un/u-n-says-it-has – credible-reports-that-china-holds-million-uighurs-in-secretcamps – idUSKBNIKVISU.

5. Mistreanu, Simina. „Life Inside Chinas Social Credit Laboratory”, *Foreign Policy*, 3 aprilie, 2018.

https://foreignpolicy.com/2018/04/03/life-inside-chinas-social-credit-laboratory/.

6. Idem.

7. „China Shames Jaywalkers Through Facial Recognition”, *Phys*.org., 20 iunie 2017. https://phys.org/news/2017 – 06-china-shames-jaywalkers-facial-recognition.html.

CAPITOLUL 1: 1 MINTEA ŞI MAŞINILE: O FOARTE SCURTĂ ISTORIE

A INTELIGENŢEI ARTIFICIALE

1. „The Seikilos Epitaph: The Oldest Song în The World”. *Wired*, *29* octombrie 2009, https://www.wired.com/2009/10/the-seikilos-epitaph.

2. „Population Clock: World”. *Census*. *Gov*., 2018,https://www.census.gov/popelock/world.

3. King, Elizabeth. „Clockwork Prayer: A Sixteenth-Century Mechanical Monk”, *Virginia Commonwealth University*, Spring 2002, http’s: //blackbird. veu.edu/vi n 1 /nonfiction/kinge/prayerintroduction.htm.

4. Hobbes, Thomas. *De Corpore Politico*, *Or*, *The Elements Of Law Moral & Politick*.

5. Descartes, René. „Meditations on First Philosophy”, Second Mediation §25, 1641. University of Connecticut.

http://selfpace. uconn.edu/class/percep/Descartes Meditations.pdf.

6. Descartes, René. *Treatise of Man*, tr. T.S. Hall. Cambridge, Mass.: Harvard University Press, 1972.

7. Leibniz, Gottfried Wilhelm. *The Monadology*, tr. Robert Latta, 1898. https://www.plato-philosophy.org/wp-content/uploads/2016/07/The-Mônadology-1714-by-Gottff ied-Wilhelm – LEIBNIZ-1646 – 1716.pdf

8. Se consideră că prima folosire a cuvântului computer” ar fi apărut într-o carte cu titlul *The Yong Mans Gleanings* scrisă de Richard Braithwaite în 1613. La acea vreme, computerele erau oamenii care efectuau calculele.

9. „Blaise Pascal”. *Biography*, https://www.biography.com/people/blaisepasca1 – 9434176.

10. Leibniz scrie în *de Progressione Dyadica*: „Acest calcul [binar] ar putea fi implementat de o maşină… cu găuri care pot fi închise şi deschise. Ele vor fi deschise în locurile care corespund unui 1 şi vor rămâne închise în locurile care corespund unui O. Prin porţile deschise mici cuburi sau pietricele pot cădea pe parcursuri, iar prin cele închise nu va putea trece nimic. Această [matrice a porţilor] se va muta de la o coloană la alta în funcţie de nevoie…!”

11. Leibniz scrie: „M-am gândit din nou la planul meu legat de un nou limbaj sau sistem de scriere pentru raţiune, care ar putea servi ca o unealtă de comunicare pentru toate naţiile… Dacă am dispune de un asemenea instrument universal, am putea discuta problemele metafizice sau etice în acelaşi mod în care putem discuta matematica sau geometria. Acesta era obiectivul meu: orice neînţelegere nu ar trebui să fie altceva decât o eroare de calcul… care să poată fi uşor corectată prin legile gramaticale ale acestui limbaj nou. Astfel, în cazul

unei discuţii controversate, doi filosofi s-ar putea aşeza la masă şi ar putea calcula, spunând, ca doi matematicieni „Hai să verificăm…”

12. „Apes to Androids: Is Man a Machine as La Mettrie Suggests?” http://www.charliemccarron.com/manamachine/.

13. Manabrea, Luigi. *Sketch of the Analytical Engine invented by Charles Babbage*. London: Richard and John E. Taylor, 1843.

14. Desmond MacHale, *The Life and Work of George Boole*: *A Prelude to the Digital Age*, New ed. (Cork University Press, 2014

15. Logicianul Martin Davis explică cel mai bine acest concept în cartea sa *The Universal Computer*: *The Road from Leibniz to Turing*: „Turing ştia că un algoritm este de regulă precizat printr-o listă de reguli pe care o persoană o poate urma într-o manieră mecanică precisă, aidoma unei reţete dintr-o carte de bucate. El a putut demonstra că o astfel de persoană ar fi limitată la câteva acţiuni extrem de simple fără să modifice rezultatul calculului. Mai apoi, dovedind că nicio maşină care ar efectua doar aceste acţiuni elementare nu ar putea să stabilească dacă o concluzie dată rezultă din premisele date…

el a putut concluziona că nu există niciun algoritm pentru această Entscheidungsproblem”.

16. Davis, Martin. *The Universal Computer*: *The Road from Leibniz to Turing Third Edition*”. Boca Raton, Fla: CRC Press, 2018.

17. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence”, Intelligence”, 31 august 1955, Stanford University, http://www-formal. stanford.edu/jme/history/dartmouth/ dartmouth.html.

18. În cadrul propunerii pentru proiectul de cercetare de vară privind inteligenţa artificială de la Dartmouth, McCarthy, Minsky, Rochester şi Shannon au propus ca următorii oameni să fie invitaţi. Am redat aici

lista originală publicată în 1955, care include şi numele companiilor şi adresele. Nu toţi invitaţii au putut participa.

Adelson, Marvin
Hughes Aircraft Company
Airport Station, Los Angeles, CA Ashby, W.R.
Bamwood House
Gloucester, England
Backus, John
IBM Corporation
590 Madison Avenue
New York, NY
Bernstein, Alex
IBM Corporation
590 Madison Avenue
New York, NY
Bigalow, J. H.
Institute for Advanced Studies
Princeton, NJ
Elias, Peter
R. L.E., MIT
Cambridge, MA
Duda, W. L.
IBM Research Laboratory
Poughkeepsie, NY
Davies, Paul M.
1317 C. 18th Street
Los Angeles, CA
Fano, R. M.
R. L.E., MIT
Cambridge, MA
Farley, B. G.
324 Park Avenue
Arlington, MA
Galanter, E. H.
University of Pennsylvania
Philadelphia, PA
Gelemter, Herbert

IBM Research Poughkeepsie, NY Glashow, Harvey A.
1102 Olivia Street
Ann Arbor, MI
Goertzal, Herbert
330 West 11th Street New York, New York Hagelbarger, D.
Bell Telephone Laboratories
Murray Hill, NJ Miller, George A.
Memorial Hall
Harvard University
Cambridge, MA
Harmon, Leon D.
Bell Telephone Laboratories
Murray Hill, NJ
Holland, John H.
E. R. I.
University of Michigan
Ann Arbor, MI
Holt, Anatol
7358 Rural Lane Philadelphia, PA
Kautz, William H.
Stanford Research Institute Menlo Park, CA
Luce, R. D.
427 West 117th Street
New York, NY
MacKay, Donald
Departament of Physics
University of London
London, WC2, England
McCarthy, John
Dartmouth College
Hanover, NH
McCulloch, Warren S.
R.L.E., M.I.T.
Cambridge, MA
Melzak, Z. A.
Mathematics Departament

University of Michigan
Ann Arbor, MI
Minsky, M. L.
112 Newbury Street
Boston, MA
More, Trenchard
Departament of Electrical Engineering
MIT
Cambridge, MA
Nash, John
Institute for Advanced Studies
Princeton, NJ
Newell, Allen
Departament of Industrial Administration Carnegie Institute of Technology
Pittsburgh, PA
Robinson, Abraham
Departament of Mathematics
University of Toronto
Toronto, Ontario, Canada
Rochester, Nathaniel
Engineering Research Laboratory
IBM Corporation
Poughkeepsie, NY
Rogers, Hartley, Jr.
Departament of Mathematics
MIT
Cambridge, MA
Rosenblith, Walter
R.L.E., M.I.T.
Cambridge, MA
Rothstein, Jérôme
21 East Bergen Place
Red Bank, NJ
Sayre, David
IBM Corporation
590 Madison Avenue

New York, NY
Schorr-Kon, J. J.
C-380 Lincoln Laboratory, MIT
Lexington, MA
Shapley, L.
Rand Corporation
1700 Main Street
Santa Monica, CA
Schutzenberger, M.P.
R.L.E., M.I.T.
Cambridge, MA
Selfridge, O. G.
Lincoln Laboratory, M.I.T.
Lexington, MA
Shannon, C. E.
R.L.E., M.I.T.
Cambridge, MA
Shapiro, Norman
Rand Corporation 1700 Main Street
Santa Monica, CA
Simon, Herbert A.
Departament of Industrial Administration Carnegie Institute of Technology Pittsburgh, PA
Solomonoff, Raymond J.
Tehnical Research Group
17 Union Square West
New York, NY
Steele, J.E., Capt. USAF
Area B., Box 8698
Wright-Patterson AFB
Ohio
Webster, Frederick
62 Coolidge Avenue Cambridge, MA
Moore, E. F.
Bell Telephone Laboratory
Murray Hill, NJ
Kemeny, John G.

Dartmouth College
Hanover, NH

19. Am compus o listă foarte scurtă de femei şi oameni de culoare care ar fi putut contribui substanţial la atelierul de la Dartmouth, dar au fost omişi. Această listă nu este exhaustivă. Aş fi putut continua pe zeci de pagini. Ea este emblematică pentru oamenii inteligenţi, capabili şi creativi care nu au fost invitaţi.

James Andrews, matematician şi profesor la Florida State University, specializat în teoria grupurilor şi teoria nodurilor.

Jean Bartik, matematiciană, unul dintre programatorii computerului ENIAC.

Albert Turner Bharucha-Reid, matematician şi teoretician cu contribuţii importante în domeniul lanţurilor Markov, teoriei probabilităţii şi statisticii.

David Blackwell, statistician şi matematician cu contribuţii importante în teoria jocului, teoria informaţiei, teoria probabilităţii şi statistica bayesiană.

Marnie Phipps Clark, doctor şi psiholog social, cu experienţă în domeniul conştiinţei de sine.

Thelma Estrin, pionier al aplicării sistemelor informatice în cercetarea neurofiziologică şi cerebrală. Cercetătoare în departamentul de electroencefalografie al Institutului de Neurologie de la Columbia Presbyterian în perioada proiectului de cercetare de vară de la Dartmouth.

Evelynn Boyd Granville, doctor în matematică, a elaborat programele folosite pentru analiza traiectoriilor în cadrul primelor misiuni cu astronauţi trimise în spaţiu şi spre lună.

Betty Holberton, matematiciană, unul dintre programatorii computerului ENIAC. Ea a inventat punctele de oprire în depanarea programelor.

Grace Hopper, informaticiană, creatoare a limbajului COBOL, folosit până în zilele noastre.

Mary Jackson, ingineră şi matematiciană, prima femeie inginer de culoare de la NASA.

Kathleen McNulty, matematiciană, unul dintre programatorii computerului ENIAC.

Marlyn Meltzer, matematiciană, unul dintre programatorii computerului ENIAC, primul computer programabil integral electronic.

Rozsa Peter, matematiciană, fondatoare a teoriei funcţiilor recursive. Frances Spence, matematiciană, unul dintre programatorii computerului ENIAC.

Ruth Teitelbaum, matematiciană, unul dintre programatorii computerului ENIAC. A calculat împreună cu Marlyn Meltzer ecuaţii pentru traiectorii balistice.

Dorothy Vaughan, matematiciană, supervizoare în 1949 a grupului West Area Computers.

Jesse Ernest Wilkins Jr., fizician nuclear, inginer mecanic şi matematician, a devenit cel mai tânăr student al universităţii din Chicago la vârsta de 13 ani.

20. „The Dartmouth Workshop – as planned and as it happened”, http://wwwformal. stanford.edu/jme/slides/dartmouth/dartmouth/node 1.html.

21. „The Dartmouth Al Archives”, http://raysolomonoif.com/dartmouth/.

22. Irving John Good, „Speculations Concerning the First Ultraintelligent Machine”, *Advances în Computers*, Volume 6 (1966): 31 – 88, https://www.sciencedirect.com/science/article/pii/S006524 580860418 O? via%3 Dihub.

23. Weizenbaum, Joseph, „ELIZA – A Computer Program for the Study of Natural Language Communication Between Man and Machine”, *Computational Lingvistics*, January, 1966, http://web. stanford.edu/class/ce 124/p36-weizenabaum.pdf

24. Stenograma completă poate fi accesată pe Github: https://github.com/codeanticode/eliza

25. „The Chicago Tribune Archives”, *Chicago Tribune*, 18 iunie 1963, http://www.chicagotribune.com/tribune-archives/.

26. H.A. Simon and A. Newell, „Heuristic Problem

Solving: The Next

Advance în Operations Research”, *Operations Research 6* (1958): 110.

27. McCarthy însuşi a dorit să lucreze cu grupul în ce priveşte ideile proprii pentru reprezentarea cunoştinţelor de bun simţ şi raţiunea, dar odată ce grupul a fost asamblat, şi-a dat seama că din grup lipseau anumiţi gânditori cheie. (în cazul lui, se aştepta să fi existat şi logicieni.)

28. „2001 Is Ne-ar but Hal Is Not”, *Los Angeles Times*, *7 ianuarie 1977*, http://articles. latimes.com/1997 – 01 – 07/news/mn-16139 lartificialintelligence

29. „Language and Machines: Computers în Translation and Linguistics”, The National Academies of Science, Engineering and Medicine, 1966, https://www.nap. edU/read/9547/chapter/l

30. Lighthill, James. „Artificial Intelligence: A Generat Survey”, *Informatics*, iulie 1972, http://www.chilton-computing.org.uk/inf/ literature/reports/lighthilireport/pOO 1.htm.

31. „Thinking Allowed, Conversations On the Leading Edge of Knowledge and Discovery, with Dr. Jeffrey Mishlove”, *Intuition.org* 1998, http://www.intuition.org/txt/minsky.htm.

32. Idem.

33. Iama inteligenţei artificiale, care a durat trei decenii, a inclus noi previziunide data aceasta sub forma unor avertismente pentru viitor. În cartea sa *Computer Power and Human Reason*, Weizenbaum susţinea că deşi inteligenţa artificială este posibilă, nu ar trebui să permitem niciodată calculatoarelor să ia decizii importante pentru că lor le vor lipsi întotdeauna calităţi umane precum compasiunea sau inteligenţa. Weizenbaum face aici o distincţie esenţială între a decide şi a alege. Decizia este o activitate legată de calcul, care poate fi programată. Alegerea este însă produsul judecăţii, nu al calculului. Capacitatea de a alege este ceea ce ne conferă în ultimă

instanţă umanitatea. Filosoful John Searle de la University of California, Berkeley susţine în articolul său „Minds, Brains and Programs” că o inteligenţă artificială generală, sau, în termenii lui, „puternică” este implauzibilă. Searle spunea că un program nu poate conferi unui computer o „minte”, o „înţelegere” sau o „conştiinţă”, indiferent cât de asemănător cu oamenii s-ar putea comporta acel program.

34. Schaeffer, Jonathan, Robert Lake, Paul Lu, Martin Bryant. „CHINOOK: The World Man-Machine Checkers Champion”.

Al Magazine, primăvara, 1966, https://www.aaai.org/ojs/index.php/ aimagazine/article/view File/1208/1109

35. Goldfarb, Ari, Daniel Trefler. „Al and International Trade”, *The National Bureau of Economic Research*, Ianuary, 2018,http://www.nber.org/papers/w24254.pdf.

36. Toby Manning, „Alpha Go”, *British Go Journal* 174 (Winter 2015 – 2016): 15, https://www.britgo.org/files/2016/deepmind/BGJ174-Alpha Go.pdf.

37. Byford, Sam. „Alpha Go retires from competitive Go after defeating world number one 3-0”, *The Verge*, May 27, 2017,https://www.theverge.com/2017/5/27/15704088/alpha go-ke-jie-game-3-resultretires-future.

38. Silver, David, Julian Schrittwieser, Karen Simonyan, et al. „Mastering the Game of Go without Human Knowledge”, Deep Mind, https://deepmind.com/documents/119/agzunformattednatu re.pdf

39. Idem.

40. Ibidem.

41. Afirmaţia a fost făcută de programatorul şef al lui Zero, David Silver, în cadrul unei conferinţe de presă.

42. Byford, Sam. „Alpha Go retires from competitive Go after defeating world number one 3-0”, *The Verge*, May

27, 2017,https://www.theverge.com/201715127115704088/alphago-ke-jie-game-3-result-retiresfuture.

43. Novet, Jordan. „Google is finding ways to make money from Alphabets Deep Mind A.I. technology”, *CNBC*, 31 martie 2018,https://www.cnbe.com/2018/03/31/how-google-makes-money-ffomalphabets-deepmind-ai-research-group.html

44. Cerejo, Roydon. „Google Duplex: Understanding the Core Technology Behind Assistants Phone Calls”, *Gadgets 360*, 10 mai 2018, http’s: //gadgets. ndtv.com/ apps/features/google-duplex-googleio-ai-google-assistant-1850326

45. Quocle, Barret Zoph”. Using Machine Learning to Explore Neural Network Architecture, *Google* Al *Blog*, 17 mai 2017, https://ai. googleblog.com/2017/05/using-machine-learning-to-explore.html.

46. Schema Winograd, propusă de informaticianul canadian Hector Levesque în 2011, prezintă o alternativă la testul Turing pentru a măsura capacităţile unei inteligenţe artificiale. Schema, pe care Levesque şi colegii au dennumit-o în onoarea informaticianului Terry Winograd de la Stanford se axează pe raţionamentele de bun simţ, provocând inteligenţa artificială să răspundă la o întrebare cu pronume ambigue. Accentul pus pe depăşirea oamenilor în competiţie directă a condus la neglijarea altor metode de a măsura şi a face inteligenţa artificială să progreseze. Schema Winograd a fost gândită ca un test multidimensional pentru că pentru a-şi putea trece e nevoie de mai mult decât de un set vast de date. Ernest Davis, Leora Morgenstern şi Charles Ortiz, trei informaticieni de la NYU, au propus o provocare a schemei Winograd, care are loc o dată pe an. Pe pagină facultăţii există un exemplu excelent: https://ce. nyu.edu/faculty/davise/papers/Winogradăchemas/WS.htm. Last accessed September 5, 2018.

Consilierii municipali au refuzat să acorde

protestatarilor autorizaţia de a demonstra pentru că aceştia [se temeau de/susţineau] violenţe. Dacă cuvântul e „se temeau” atunci aceştia” se referă la consilieri. Dacă e susţineau” atunci aceştia se referă la protestatari.

În cadrul articolului, LEvesque afirmă că schemele Winograd ar trebui să respecte următoarele constrângeri:

Să poată fi uşor dezlegate de un cititor uman (în mod ideal, atât de uşor încât cititorul să nici nu sesizeze prezenţa vreunei ambiguităţi.

Să nu poată fi rezolvate prin tehnici simple cum ar fi restricţiile de selecţie

Să fie imune la Google; cu alte cuvinte, nu există un test evident statistic aplicat bazelor de date cu text care să poată dezlega frazele corect

CAPITOLUL 2: LUMEA INSULARĂ A TRIBURILOR INTELIGENŢEI ARTIFICIALE

1. Mike Isaac şi Sheera Frenkel, „Facebook Security Breach Exposes Accounts of 50 Million Users”, *New York Times*, 28 septembrie 2018, https://www.nytimes.com/2018/09/28/technology/facebook-hackdata – breach.html.

2. Casey Newton, „Facebook Portals Claims to Protect User Privacy Are Falling Apart, *The Verge*, 17 octombrie 2018, https://www

theverge.com/2018/10/17/17986992/facebook-portal-privacy – claims-ad-targeting.

3. „AMA: We are the Google Brain te-am. Wed love to answer your questions about machine learning”, *Reddit*, 2016, https://www.reddit.com/r/Machinelearning/comments/4 wbtsv/ama wearethegooglebrainteamwedloveto/.

4. Idem.

5. „Diversity”, *Google*, https://diversity. google/.

6. Tiku, Nitasha. „Googles Diversity Stats Are Still Very Dismal”, Wired, 14 august 2018, https://www.wired.com/story/googlesemployee-diversity-numbers-havent-really-improved/.

7. Daisuke Wakabayashi and Katie Benner, „How Google Protected Andy Rubin, the Father of Android, „*New York Times*, 25 octombrie 2018, https://www.nytimes.com/2018/10/25/technology/googlesexual – harassment-andy-rubin.html.

8. Broockman, David, Greg F. Ferenstein, Neil Malhotra. „The Political Behavior of Wealthy Americans: Evidence from Technology Entrepreneurs”, *Stanford Graduate School of Business*, 9 decembrie 2017, https://www.gsb. stanford.edu/faculty-research/working-papers/ political-behaviorwealthy-americans-evidence-technology.

9. „ICYMI: RNC Chairwoman & Brad Parscale Demand Answers from Facebook & Twitter”, *Republican National Committee*, *24* mai 2018, https://www.gop.comzicymime-chairwoman-brad-parscaledemand-answers-from-facebook-twitter.

10. Conger, Kate, Sheera Frenkel. „Dozens at Facebook Unite to Challenge Its Intolerant’ Liberal Culture” *New York Times*, 28 august 2018, https://www.nytimes.com/2018/08/28/technology/inside-facebook-employees-political-bias.html.

11. Rocha, Veronica. „Crime-fighting robot hits, rolls over child at Silicon Valley mall”, *Los Angeles Times*, 14 iulie 2016,http://www.latimes.com/local/lanow/lame-ln-crimefighting-robothurts-child-bay-area-20160713-snap-story.html.

12. Benson, Julian. „Elites Al Created Super Weapons and Started Hunting Players. Skynet is Here”, *Kotaku*, 3 iunie 2016,http://www.kotaku.co.uk/2016/06/03/elites-ai-created-super-weapons-and-started-h unting-players-skynet-is-here.

13. Bucks County Courier Times Newspaper Archives, 20 august 1974, p. 9

14. Există un articol grozav despre Pryor înainte să devină celebru. McPherson, James. „The new comic style of Richard Pryor”, *New York Times*, April 27, 1975

15. Vance, Ashlee. „How We Got Here”, *Bloomberg Businessweek*, *21* mai 2018, https://www.scribd.com/article/379513106/How-We-Got-Here.

16. „Computer Science”, *Stanford University*, 2018 – 2019,https://exploredegrees. stanford.edu/schoolofengineering/computerscience/e bachelortext.

17. „Vector Representations of Words”, *Tensor Flow*, https://www.tensorflow.org/tutorials/representation/worzec .

18. Bolukbaşi, Tolga, Kai-Wei Chang, et al., „Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings”, *Cornell University Library*, 21 iulie 2016, https://arxiv.org/abs/1607.06520.

19. Saltiel, Natalie. „The Ethics and Governance of Artificial Intelligence”, *MIT Media Lab*, 16 noiembrie 2017, https://www.media. mit.edu/ courses/the-ethics-and-governance-of-artificial-intelligence/.

20. Puteţi urmări cursurile aici: https://www.media. mit.edu/courses/ the-ethics-and-governance-of-artificial-intelligence/.

21. Ashcraft, Catherine, Brad McLain, Elizabeth Eger. „Women în Tech: The Facts”, *National Center for Women & Inform*ation *Technology*, 2016, https://www.newit.org/sites/default/files/resources/ womenintechfactsfulireport05132016.pdf.

22. „Degrees în computer and information sciences conferred by degree-granting institutions, by level of degree and sex of student:

1970 – 71 through 2010 – 11”, *National Center for Education Statistics*, https://nees. ed.gov/programs/digest/dl2/tables/dtl2349.asp.

11. Doctors degrees conferred by postsecondary institutions, by race/ ethnicity and field of study: 2013 – 14 and 2014 – 15”, *National Center for Education Statistics*, https://nees. ed.gov/programs/digest/d 16/tables/dt

16324.25.asp? current = yes.

12. Mims, Christopher. „What the Google Controversy Misses: The Business Case for Diversity”, *Wall Street Journal*, 13 august 2017, http’s: //www.wsi.com/articles/what-the-google-controversy-missesthe-business-case-for-diversity-1502625603.

13. Hempel, Jessi. „Melinda Gates and Fei-Fei Li Want to Liberate Al from „Guys With Hoodies”, *Wired*, 4 mai 2017,https://www.wired.com/2017/05/melinda-gates-and-fei-fei-liwant-to-liberate-ai-from-guys-with-hoodies/.

14. Meng Jing. „China looks to school kids to win the global Al race”, *South China Morning Post*, *International Edition*, 3 mai 2018, http’s: //www.semp.com/tech/china-tech/article/2144396/china-looksschool-kids-win-global-ai-race.

15. „China launches first university program to traininti Al talents”, *Zhongguancun Science Park*, 4 aprilie 2018,http://www.chinadaily.com.cn/m/Beijing/zhongguancun/

2018 – 04/04/content35979394.htm.

16. Barboza, David. „The rise of Baidu (That’s Chinese for Google)”, *New York Times*, 1 septembrie 2006,https://www.nytimes.com/2006/09/17/business/yourmoney/17 baidu.html.

17. „Râse Of Chinas Big Tech în Al: What Baidu, Alibaba, And Tencent Are Working On”, *CB Insights*, 26 aprilie 2018,https://www.chinsights.com/research/china-baidu-alibaba-tencent-artificial-intelligence-dominance/.

18. Lucas, Louise. „The Chinese Communist party entangles big tech”, *Financial Times*, 18 iulie 2018, https://www.fi.com/content/5 dOafăc4 – 846 c-l Ie8-a29 d-73 eăd454535 d.

19. „A Hong Kong Newspaper on a Mission to Promote Chinas Soft

Power”, *New York Times*, 31 martie 2018,https://www.nytimes.com/2018/03/31/world/asta/south-chinamorning-post-hong-kong-alibaba.html.

20. Farhi, Paul”. Washington Post closes sale to Amazon founder Jeff Bezos”, *Washington Post*, 1 octombrie 2013,https://www.Washingtonpost.com/business/economy/ Washington-post-closes-sale-to-amazon-founder-jeff-bezos/2013/10/01/ea Abiba-2 acf-lle3 – 97 a3-ff27582 28523 story.html? noredirect = on & utmterm =.3 d04830 eab75

21. Lim, Jason. „WeChat Is Being Trialled To Make Hospitals More Efficient în China”, *Forbes*, 16 iunie 2014 https://www.forbes.com/sites/jlim/2014/06/16/wechat-is-beingtrialed-to-make-hospitals-more-efficient-in-china/63 zd315 se2

22. „Râse Of Chinas Big Tech în Al: What Baidu, Alibaba, And Tencent Are Working On”, *CB Insights*, *26* aprilie 2018,https://www.chinsights.com/research/china-baidu-alibaba-tencent-artificial-intelligence-dominance/.

23. Kharpal, Arjun. „Chinas Tencent surpasses Facebook în valuation a day after breaking $500 billion barrier”, *CNBC*, 21 noiembrie 2017, https://www.cnbe.com/2017/11/21 /tencent-surpasses-facebookin-valuation.html.

24. Rutherford, Sam. „5 Things to Know About Tencent, the Chinese Internet Giant That’s Worth More Than Facebook Now”, *Gizmodo*, November 27, 2017, https://gizmodo.com/5-things-to-know-abouttencent-the-chinese-internet-gi-1820767339.

25. Fannin, Rebecca. „China Releases a Tech Dragon: the BAT”, *Techonomy*, 23 mai 2018, https://techonomy.com/2018/05/ china-releases-tech-dragon-bat/.

26. „Mobile Fact Sheet”, *Pew Research Center*, 5 februarie 2018,http://www.pewinternet.org/fact-sheet/mobile/.

27. Yurieff, Kaya. „Amazons Cyber Monday was its biggest sales day ever”, *CNN Money*, 29 noiembrie 2017, https://money. cnn.com/2017/11/29/technology/amazon-cybermonday/index.html.

28. Wang, Helen H. „Alibaba’s Singles’ Day By The Numbers: A Record $25 Billion Haul”, *Forbes*, 12 noiembrie 2017.

29. Fannin, Rebecca. „China Releases a Tech Dragon: the BAT”, *Techonomy*, 23 mai 2018, https://techonomy.com/2018/05/chinareleases-tech-dragon-bat/.

30. Brown, Michael, Pavneet Singh. „China’s Technology Transfer Strategy”, *Defense Inovation Unit Experimental*, februarie 2017, https://new. reorg-research.com/data/documents/20170928/ 59 ccf7 de70 z.pdf.

31. Pentru cei care doresc să acceseze textul integral al celui de-al 13-lea plan cincinal, el poate fi consultat la adresa http://www.gov.cn/ xinwen/2016 – 03/17 /content5054992.htm. 17 martie 2016, traducere în limba engleză.

32. J.P. „What is Chinas belt and road inițiative?”, *The Economist*, 15 mai 2017, https://www.economist.com/the-economistexplains/2017/05/ 14/what-is-chinas-belt-and-road-initiative.

33. Babones, Salvatore. „China’s Middle Class Is Pulling Up the Ladder Behind Itself?” *Foreign Policy*, 1 februarie 2018,https://foreignpolicy.com/2018/02/01/chinas-middle-class-ispulling-up-the-ladder-behind-itself/.

34. „The American Middle Class Is Losing Ground”, *Pew Research Center*, 9 decembrie 2015,http://www.pewsocialtrends.org/2015/12/09/theamerican-middle-class-islosing-ground/.

35. Martin, Emmie. „70% of Americans consider themselves middle class-but only 50% are”, *CNBC*, 30 iunie 2017,https://www.cnbe.com/2017/06/30/70-percent-of-americansconsider-themselves-middle-class-but-only-50-percent-are.html

36. Bhattarai, Abha. „China asked Marriott to shut down its website. The company complied”, *Washington*

Post, 18 ianuarie 2018.

37. Jacobson, Louis. „Yes, Donald Trump did call climate change a Chinese hoax”, *Politi Fact*, 3 iunie 2016, https://www.politifact.com/trutho-meter/statements/2016/jun/03/hillary-clinton/yes-donald-trump-di d-call-climate-change-chinese-h/.

38. Greenstone, Michael. „Four Years After Declaring War on Pollution, China Is Wining”, *New York Times*, 12 martie 2018,https://www.nytimes.com/2018/03/12/upshot/china-pollution-environment-longerlives.html https://www.nytimes.com/2018/03/12/upshot/china-pollutionenvironment-longer-lives.html.

39. Fordham, Carl Gene. „20 Actually Useful Chengyu”, *Cangene.com*, 14 August 2008, http://cangene.com/blog/2010/07/20-actually-usefulchengyu-%E6%88%90%E8%AF%AD.

40. Chen, Stephen. „China takes surveillance to new heights with flock of robotic Doves, but de they come în peace?” *South China Morning Post*, 24 iunie 2018, https://www.semp.com/news/China/society/ article/2152027/china-takes-surveillance-new-heights-flock-roboticdoves-do-they

41. Stewart, Phil. „China racing for Al military edge over U.S.: report”, *Reuters*, 27 noiembrie 2017, https://www.reuters.com/article/us-usachina-ai/china-racing-for-ai-military-edge-over-u-s-reportid USKBNIDSOG5.

42. Conger, Kate. „Google Employees Resign în Protest Against Pentagon Contract”, *Gizmodo*, 14 mai 2018, https://gizmodo.com/googleemployees-resign-in-protest-against-pentagon-con-1825729300.

43. Nitasha Tiku, Amazons Jeff Bezos Says Tech Companies Should Work with the Pentagon”, *Wired*, October 15, 2018.

https://www.wired.com /story/amazons-jeff-bezos-says-techeompanies-should-work-with-the – pentagon/.

44. Stewart, Phil. „China racing for Al military edge over U.S.: report”, *Reuters*, 27 noiembrie 2017, https://www.reuters.com/article/us-usachina-ai/china-racing-for-ai-military-edge-over-u-s-reportid USKBNIDSOG5.

7. „China issues guideline on artificial intelligence development”, *The State Council*, *Peoples Republic of China*, 20 iulie 2017, http://english.gov.cn/policies/latestreleases/2017/07/20/content281 475742458322.htm.

58. „Key Al guidelines unveiled”, *The State Council*, *Peoples Republic of China*, 15 decembrie 2017, http://english.gov.cn/statecouncil/ministries/2017/12/15/content281475977265006.htm.

59. Kania, Elsa B. „Chinas Al Giants Can’t Say No to the Party”, *Foreign Policy*, *2* august 2018, https://foreignpolicy.com/2018/08/02/chinasaigiants-cant-say-no-to-the-party/.

60. Idem.

61. Ibidem.

62. Pomfret, John. „Chinas new surveillance state puts Facebooks privacy problems in the shade”, *Washington Post*, 27 martie 2018, https://www.Washingtonpost.com/news/globalopinions/wp/2018/03/27/chinas-new-surveillance-state-puts-facebooks-privacy-problems-in-the-shade/? noredirect = on & utmterm =. (8 ccOacd08 bl.

63. Wright, Nicholas. „How Artificial Intelligence Will Reshape the Global Order”, *Foreign Affairs*, 10 iulie 2018,https://www.foreignaffairs.com/articles/world/2018 – 07 – 10/howartificial-intelligence-will-reshape-global-order.

64. Zhang Hongpei. „Many netizens take issue with Baidu CEOS comments on data privacy” *Global Times*, 26 martie 2018,http://www.globaltimes.cn/content/1095288.shtml

65. Zhong, Raymond. „Chinese Tech Giant on Brink of Collapse în New U.S. Cold War”, *New York Times*, 9 mai

2018

66. Sacks, Samm. „Beijing Wants to Rewrite the Rules of the Internet”, *The Atlantic*, 19 iunie 2018, https://www.theatlantic.com/international/archive/2018/06/zte-huawei-china-trump-trade-cyber/563033/.

67. Idem.

68. Ibidem.

69. „The Recruitment Program for Inovative Talents (Long Term)”, *The Thousand Talents Plan*, http://1.000 plan.org/en/.

70. Marquand, Michael. „The Trump Administration Plays Catch-Up on Artificial Intelligence”, *Wired*, 11 mai 2018,https://www.wired.com/story/trump-administraţion-plays-catch-upartificial-intelligence/.

71. Levy, Ari. „Dropbox is going public: Here’s who’s making money”, *CNBC*, 23 februarie 2018, https://www.cnbe.com/2018/02/23/ dropbox-is-going-public-heres-whos-making-money.html.

72. Gramlich, John. „5 facts about Americans and Facebook”, *Pew Research Center*, 10 aprilie 2018, http://www.pewresearch.org/ fact-tank/2018/04/10/5-facts-about-americans-and-facebook/

73. Weise, Elizabeth. „Amazon Prime is popular, but în three-quarters of all U.S. homes? That’s open to debate”, *USA Today*, 20 octombrie 2017, https://www.usatoday.com/story/tech/2017/10/20/amazon-primebig-though-how-big-no-one-knows/784695001/.

74. „Mobile Fact Sheet”, *Pew Research Center*, 5 februarie 2018,http://www.pewinternet.org/fact-sheet/mobile/.

75. *Tensorflow*, https://github.com/tensorflow/tensorflow.

76. „Microsoft to acquire GitHub for $7.5 billion”, *Microsoft News Center*, 4 iunie 2018, https://news. microsoft.com/2018/06/04/microsoft-toacquire-github-for-7 – 5-billion/.

77. Novet, Jordan. „Why tech companies are racing

ea ch other to make their own custom A.I. chips”, *CNBC*, 21 aprilie 2018,https://www.cnbe.com/2018/04/21/alibaba-joins-google-others-inmaking-custom-ai-chips.html.

78. The full paper can be accessed at https://graphics. axios.com/pdf *I* Platform Policy Paper.pdf?ga = 2.167458877.2075880604.1541172609 – 1964512884.1536872317.

79. Aceste mesaje pot fi accesate la adresa https://twitter.com/timcook/status /1055035534769340418.

CAPITOLUL 3: O MIE DE TĂIETURI CU HÂRTIA: CONSECINŢELE NEINTENŢIONATE ALE INTELIGENŢEI ARTIFICIALE

1. „An Owners’ Manual for Google Shareholders”, *Alphabet Investor Relations*, https://abc. xyz/investor/founders-letters/2004/ipo-letter.html.

2. Idem.

3. „Leadership Principles”, *Amazon Jobs*, https://www.amazon. jobs/ principles.

4. „Focus on Impact”, Facebook, 8 septembrie 2015,https://www.facebook.com/facebookcareers/photos/a.1655178611435493.1073741828.1633466236940064/1655179928102028/? type = 3 & theater.

5. „Core Values”, Tencent, https://www.tencent.com/en-us/culture.html.

6. „Culture and Values”, Alibaba Group, https://www.alibabagroup.com/en/about/culture.

7. Bergen, Mark. „Google Engineers Refused to Build Security Tool to Win Military Contracts”, *Bloomberg*, 21 iunie 2018,https://www.bloomberg.com/news/articles/2018 – 06 – 21/googleengineers-refused-tobuild-security-tool-to-win-military-contracts.

8. Pichai, Sundar. „Al at Google: our principles”, *Google*, 7 iunie 2018, https://www.blog. google/technology/ai/ai-principles/.

9. „Quick Facts, *U.S. Census Bureau*, 1 iulie 2017,https://www.census.gov/quickfacts/fact/table/US/PST045217.

10. McCormack, Alan, John Rusnak, Carliss Baldwin. „Exploring the Duality between Product and Organizaţional Architectures: A Test of the „Mirroring” Hypothesis”, *Harvard Business School*, https://www.hbs.edu/faculty/Publication%20 Files/08 – 039186 le507 – 1 de1 – 4602 – 85 b8 – 90 d71559 d85 b.pdf.

11. Miotto, Riccardo, Brian A. Kidd, Joel T. Dudley. „Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records”, *Scientific Reports*, 17 mai 2016, https://www.nature.com/articles/step26094.

12. Mordvintsev, Alexander, Christopher Olan, Mike Tyka. „Începţionism: Going Deeper into Neural Networks”, *Google* Al *Blog*, June 17, 2015, https://ai. googleblog.com/2015/06/începţionism-goingdeeper-into-neural.html.

13. „Începţionism: Going deeper into Neural Networks”, *Google Photos*, 17 iunie 2015, https://photos. google.com/share/AFIQipPXOSC170 z-Wilt9 LnuoliattX4OUCi8EP65 ci Vnbmă 1 my gsGQAieQUC 1VQW dgQ? key = avBxWihw Szg2 RjJWLWRuvFBBZENld205 budEMnhb.

14. Sweeney, Latanya”. Discrimination in Online Ad Delivery”, *Harvard University*, 28 ianuarie 2013, https://arxiv.org/abs/1301.6822.

15. Winston, Ali. „Palantir Has Secretly Been Using New Orleans to Test Its Predictive Policing Technology”, *The Verge*, 27 februarie 2018, https://www.theverge.com/2018/2/27/17054740/palantir-predictivepolicing-tool-new-orleans-nopd.

16. Angwin, Julia, Jeff Larson, Surya Mattu, Lauren Kirchner. „Machine Bias”, *ProPublica*, 23 mai 2016, https://www.propublica.org/article/machine-bias-risk-assessmentsin-criminal-sentencing.

17. McLaughlin, Kevin, Jessica E. Lessin. „Deep Confusion: Tensions Lingered Within Google Over Deep

Mind”, *The Information*, 19 aprilie 2018.

18. Vincent, James. „Google’s Deep Mind and UK hospitals made illegal deal for health data, says watchdog”, *The Verge*, 3 iulie 2017, http’s: //www.theverge.com/2017/7/3/15900670/google-deepmindroyal-fr ce-2015-data-deal-ico-ruling-illegal.

19. Suleyman, Mustafa, Dominic King”. The Information Commissioner, the Royal Free, and what we ve learned”, *Deep Mind* blog, 3 iulie 2017, https://deepmind.com/blogzico-royal-ffee/.

20. „Microsoft Launches Fifth Generation of Popular Al Xiaoice”, *Microsoft News Center*, https://www.microsoft.com/enus/ard/news/newsinfo.aspx? newsid = article2017091.

21. Kleeman, Sophie”. Here Are the Microsoft Twitter Bot’s Craziest Racist Rants”, *Gizmodo*, 24 martie 2016, https://gizmodo.com/ here-are-the-microsoft-twitter-bot-s-craziest-racist-ra-1766820160.

22. Lee, Peter. „Learning from Tay’s introduction, 25 martie 2016, *Microsoft Official Blog* https://blogs. microsoft.com/blog/2016/03/25/learning-tays-introduction/.

23. Harding, Verity and Sean Legassick. „Why we launched Deep Mind Ethics & Society”, *Deep Mind blog*, https://deepmind.com/blog/whywe-launched-deepmind-ethics-society/.

24. „Chinese hate us, Baidu CEO admits as he vows people before profit, *Reuters*, 10 mai 2016, https://www.cnbe.com/2016/05/10/baidu-ceotells-staff-to-put-values-before-profit-after-cancer-deathscandal.html.

CAPITOLUL 4 DE AICI LA SUPER-INTELIGENŢA ARTIFICIALĂ: SEMNALELE DE AVERTIZARE

1. Am modelat scenariile din partea a doua folosind studii din diferite surse. Referinţele se găsesc în bibliografie. De asemenea, am vizitat expoziţia *Robots* de la Science Museum (Londra) care explorează ultimii 500 de ani de roboţi umanoizi, un loc minunat în care am putut explora temele prezentate în capitolele 5 – 7.

2. Floorwalker, Mike. „10 Deadly Disasters We Should Have Seen Coming”, *Listrerse*, *2* martie 2013, https://listrerse.com/2013/03/02/10 – deadly-disasters-we-should-have-seen-coming/. Vezi şi Teather. David. „90-second nightmare of shuttle crew”, *The Guardian*, 6 februarie 2003, https://www.theguardian.com/world/2003/feb/06/Columbia. science.

3. Brooker, Katrina. „I Was Devastated: Tim Berners-Lee, the Man Who Created to World Wide Web, Has Some Regrets”, *Vanity Fair*, august 2018, https://www.vanityfair.com/news/2018/07/the-man-who-created-the-worid-wide-web-has-some-regrets.

4. Berners-Lee, Tim. „The web is under threat. Join us and fight for it.”, *World Wide Web Foundation*, 12 martie 2018,https://webfoundation.org/2018/03/web-birthday-29/.

5. „Subscriber share held by smartphone operating sistems in the United States from 2012 to 2018”, *Statista*, https://www.statista.com/statistics/266572/market-share-held-by-smartphone-platforms-in-the-unit ed-states/.

6. „Primary e-mail providers according to consumers in the United States as of 2016, by age group”, *Statista*, https://www.statista.com/statistics/547531/e-mail-provider-ranking-consumer-usa-age/.

7. Fernandez, Marisa. „Amazon leaves retail competitors in the dust, claims 50% of U.S. e-commerce market”, *Axios*, 13 iulie 2018,https://www.axios.com/amazon-now-has-nearly-50-of-the-us-e-commerce-market-1531510098 – 8529045 a-508 d-46 d6 – 86 a

1 dOz4 a04 b4.html.

8. Kleiner, Art. „The Man Who Saw the Future”, *Strategy + Business*, 12 februarie 2003, https://www.strategybusiness.com/article/8220? gko = Od07 f.

9. Cass R. Sunstein, „Probability Neglect: Emotions,

Worst Cases, and Law”, Chicago Unbound, John M. Olin Program în Law and Economics Working Paper No. 138, 2001.

10. „Quick Facts 2015”, *National Highway Traffic Safety Administration*, https://crashstats. nhtsa. dot.gOv/Api/Public/View Publication/812348.

11. „Aviation Statistics”, *National Transportation Safety Board*, https://www.ntsb.gov/investigations/data/Pages/aviations tats. aspx.

12. Brooks, Frederick P. *The Mythical Man Month*: *Essays on Software Engineering Anniversary Edition*. Addison Wesley; 2 d edition.

2 august 1995.

13. Wilby, Peter. „Beyond the Flynn efect: new myths about race, family and IQ?” *The Guardian*, 27 septembrie 2016,https://www.theguardian.com/education/2016/sep/27/ james-flynn-race-iq-myths-doesyour-family-make-you-smarter.

14. Condon, Stephanie. „US orice again boasts the worlds fastest supercomputer”, *ZONet*, 8 iunie 2018, https://www.zonet.com/article/usonce-again-boasts-the-worids-fastest-supercomputer/.

15. Viegas, Jen. „Comparison of Primate Brains Reveals Why Humans Are Unique”, *Seeker*, 23 noiembrie 2017, https://www.seeker.com/health/mind/comparison-of-primate-brainsteveals-why-humans-are-unique.

16. Bostrom, Nick. „Ethical Issues în Advanced Artificial Intelligence”, *nickbostrom.com*, 2003, https://nickbostrom.com/ethics/ai.html.

17.1. J. Good, „Speculations Concerning the First Ultraintelligent Machine”, *Advances în Computers* 6 (1965): 31 – 88.

18. Pratt, Gill A. „Is a Cambrian Explozion Coming for Robotics?”*American Economic Association*, vara 2015,https://www.aeaweb.org/articles? id = 10.1257/jep.29.3.51.

CAPITOLUL 6 CUM SĂ ÎNVĂŢĂM SĂ TRĂIM CU MILIOANE DE TĂIETURI CU HÂRTIA: SCENARIUL PRAGMATIC

1. Ross, Casey. „IBM Watson Health hampered by internal rivalries and disorganization, former employees say”, *STAT*, 14 iunie 2018, https://www.statnews.com/2018/06/ 14/ibm-watson-health-rivalriesdisorganization/.

2. Idem.

3. Elsayed, Gamaleldin, Ian Goodfellow and Jascha Sohl Dickson. „Adversarial Reprogramming of Neural Networks”, preprint edition accessed at https://arxiv. Org/pdf/1806. 1 1146.pdf.

4. Wang, Orange. „Chinese mobile payment giants Alipay, Tenpay fined USŞ88.000 for breaking foreign exchange rules”, *South China Morning Post*, 25 iulie 2018, https://www.semp.com/news/China/economy/ article/2156858/chinese-mobile-paymentgiants-alipay-tenpayfined-us88.000.

CAPITOLUL 7 DINASTIA RENGONG ZHINENG: SCENARIUL CATASTROFIC

1. „China Has a Vastly Ambitious Plan to Connect the World”, *The Ec onomist*, 28 iulie 2018, https://www.economist.com/briefing/2018/07/26/china-has-a-vastlyambitious-plan-to-connect-the-world.

2. Idem.

3. Ibidem.

4. Ibidem.

5. Londono, Ernesto. „From a Space Station in Argentina, China Expands Its Reach în Latin America”, *New York Times*, 28 iulie 2018, http’s: //www.nytimes.com/2018/07/28/world/americas/china-latinamerica.html.

6. Kochanek, Kenneth D, Sherry L. Murphy, Jiaquan Xu, Elizabeth Areas. „Mortality in the United States, 2016”, *NCHS Data Brief*, https://www.cde.gov/nehs/data/databriefs/db293.pdf

CAPITOLUL 8 PIETRICELE ŞI BOLOVANI: CUM SĂ REPARĂM VIITORUL INTELIGENŢEI ARTIFICIALE

1. „Vinton G. Cerfi *Google AI*, https://ai. google/research/people/ author32412.

2. Povestirea Runaround (Fuga în cerc) a fost publicată pentru prima dată în numărul din martie 1942 al revistei *Astounding Science Fiction*. Ea face parte şi din colecţiile de povestiri *I, Robot* (*1950*), *The Complete Robot* (*1982*) *şi Robot Visions* (*1990*).

3. „Human Cell Atlas”, https://www.humancellatlas. Org/learn-mOre.

4. Metz, Cade. „As China Marches Forward on A.I., the White House Is Silent”, *New York Times*, 12 februarie 2018,https://www.nytimes.com/2018/02/12/technology/chin a-trumpartificial-intelligence.html.

5. Heisler, Yoni. „Amazon în 2017 spent almost twice as much on R & D as Microsoft and Apple – combined”, *BGR*, 10 aprilie 2008, http’s: //bgr.com/2018/04/10/amazon-vs-apple-research-and-development-2017-alphabet-google/

6. „The OTA Legacy”, *Princeton University*, http://www.princeton.edu/-ota/.

7. „Dining”, *Departament of Defense*, http://www.whs.mil/ our-services/building-facilities/dining.

8. „The Spheres”, https://www.seattlespheres.com/.

9. Adamczyk, Alicia. „These Are the Companies With the Best Parental Leave Policies”, *Money*, 4 noiembrie 2015, http://time.com/money/4098469/paid-parental-leave-googleamazon-apple-facebook/.

10. Webb, Amy. „Apple vs. FBI debate may be the least of our challenges”, *CNN*, 29 februarie 2016 https://www.cnn.com/2016/02/25/opinions/when-technology-clashes-with-law-iphone-opinionwebb/ index.html.

11. „China Un covers 500.000 Food Safety Violations in Nine Months”, *Reuters*, 24 decembrie 2016, https://www.reuters.com/article/us-china – food-

safety/china-uncovers-500000-food-safety-violations-in-nine – months-idUSKBN 14D046.

12. Tandon, Suneera. „An IBM te-am identified deep gender bias from 50 years of Booker Prize shortlists”, *Quartz India*, 24 iulie 2018, http’s: //qz.com/India/1333644/ibm-identifies-gender-bias-in-bookerprize-novel-shortlists/.

13. Mason, Hillary. *Twitter*, 28 martie 2018, https://twitter.com/hmason/status/979044821749895170